

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**



①9 **BUNDESREPUBLIK  
DEUTSCHLAND**



**DEUTSCHES  
PATENT- UND  
MARKENAMT**

⑫ **Offenlegungsschrift**  
⑩ **DE 101 13 781 A 1**

⑤① Int. Cl.<sup>7</sup>:  
**C 07 K 14/435**

②① Aktenzeichen: 101 13 781.8  
②② Anmeldetag: 21. 3. 2001  
④③ Offenlegungstag: 13. 12. 2001

**DE 101 13 781 A 1**

⑥⑥ Innere Priorität:

100 28 212. 1      09. 06. 2000  
100 53 478. 3      24. 10. 2000

⑦① Anmelder:

IPK-Institut für Pflanzengenetik und  
Kulturpflanzenforschung, 06466 Gatersleben, DE

⑦④ Vertreter:

Maiwald Patentanwalts-GmbH, 80335 München

⑦② Erfinder:

Scheller, Jürgen, Dr., 06484 Quedlinburg, DE;  
Conrad, Udo, Dr., 06458 Hausneindorf, DE; Große,  
Frank, Prof. Dr., 07743 Jena, DE; Gührs, Karl-Heinz,  
Dr., 07745 Jena, DE

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen**

⑤④ Synthetische Spinnenseidenproteine und deren Expression in transgenen Pflanzen

⑤⑦ Die Erfindung betrifft eine DNA-Sequenz, die für ein synthetisches Spinnenseidenprotein kodiert, rekombinante Spinnenseidenproteine, die durch die erfindungsgemäße DNA-Sequenz kodiert sind, Verfahren zur Herstellung von Pflanzen bzw. Pflanzenzellen, die rekombinantes Spinnenseidenprotein enthalten sowie transgene Pflanzenzellen und Pflanzen, die eine DNA-Sequenz enthalten, die für ein synthetisches Spinnenseidenprotein kodiert. Des weiteren betrifft die Erfindung ein Verfahren zur Gewinnung von pflanzlichem Spinnenseidenprotein aus transgenen Pflanzen sowie pflanzliche Spinnenseidenproteine, die nach einem derartigen Verfahren hergestellt worden sind.

**DE 101 13 781 A 1**

- [0001]** Die Erfindung betrifft eine DNA-Sequenz, die für ein synthetisches Spinnenseidenprotein kodiert, rekombinante Spinnenseidenproteine, die durch die erfindungsgemäße DNA-Sequenz kodiert sind, Verfahren zur Herstellung von Pflanzen bzw. Pflanzenzellen, die rekombinantes Spinnenseidenprotein enthalten sowie transgene Pflanzenzellen und Pflanzen, die eine DNA-Sequenz enthalten, die für ein synthetisches Spinnenseidenprotein kodiert. Des weiteren betrifft die Erfindung ein Verfahren zur Gewinnung von pflanzlichem Spinnenseidenprotein aus transgenen Pflanzen sowie pflanzliche Spinnenseidenproteine, die nach einem derartigen Verfahren hergestellt worden sind.  
RN : MH : uh
- [0002]** Spinnenseide weist hervorragende mechanische Eigenschaften auf, die jene vieler bekannter natürlicher und künstlicher Materialien übertrifft. Hauptbestandteile der Spinnenseide sind Faserproteine wie beispielsweise Fibroin aus dem Seidenspinner sowie Spidroin 1 und Spidroin 2 aus *Nephila clavipes*. Die Festigkeit und Elastizität der Seide beruht auf der Gegenwart von kurzen repetitiven Aminosäure-Einheiten, die in diesen natürlichen Proteinen vorliegen. Diese mechanischen Eigenschaften prädestinieren die Spinnenseide für eine Reihe von verschiedensten technischen Anwendungen wie beispielsweise die Herstellung von stabilen Fäden bzw. Seiden. Ferner verfügen die Spinnenseidenfäden aufgrund ihrer proteinchemischen Eigenschaften über ein geringes immunogenes und allergenes Potential, weshalb sich in Kombination mit den mechanischen Eigenschaften eine Anwendung in der Medizin beispielsweise als natürliches Garn zum Vernähen von Wunden, als Anheftungsflächen für kultivierte Zellen, als Gerüste für künstliche Organe und dergleichen anbietet.
- [0003]** Voraussetzung für eine derartige technische bzw. medizinische Nutzung der Spinnenseide ist jedoch die Herstellung von Spinnenfäden bzw. Spinnenseidenproteinen in großem Maßstab. Zu diesem Zweck wurde bislang versucht, die für die Produktion der Spinnenseide verantwortlichen Spidroin- bzw. Fibroingene in *E. coli* zu exprimieren. Die sich häufig wiederholenden Sequenzen in den entsprechenden Genen gehen jedoch bei der Reproduktion in Bakterien nach und nach verloren. Ein weiteres Problem ist die Größe der genetischen Information, die für das Bakterium zu umfangreich zu sein scheint, so daß die Spinnenseiden-Gene nicht immer vollständig ausgelesen werden.
- [0004]** Versuche der Expression in Hefezellen ergaben zwar stabilere und längere Seidenproteine, die Fäden, die daraus gesponnen wurden, weisen jedoch nicht die selben vorteilhaften Eigenschaften der natürlichen Seide auf, so daß beispielsweise eine medizinische Anwendung einer derart synthetisch hergestellten Seide nicht möglich ist. Es besteht somit ein Bedarf an synthetischen Seidenproteinen, die in technischem Maßstab hergestellt werden können und nach dem Verspinnen zu Fäden mechanische Eigenschaften aufweisen, die mit jenen der natürlichen Seide vergleichbar sind.
- [0005]** Aufgabe der vorliegenden Erfindung ist es deshalb, DNA-Sequenzen bereitzustellen, die für ein synthetisches Spinnenseidenprotein kodieren, das eine möglichst große Ähnlichkeit mit den bisher bekannten natürlichen Sequenzen von Faserproteinen der Spinnenseide aufweist. Ferner ist es Aufgabe der vorliegenden Erfindung, ein Verfahren bereitzustellen, mit dem synthetische Spinnenseidenproteine in großem Maßstab hergestellt werden können.
- [0006]** Weitere Aufgaben der vorliegenden Erfindung ergeben sich aus der folgenden Beschreibung.
- [0007]** Oben genannte Aufgaben werden durch die Merkmale der unabhängigen Schutzansprüche gelöst.
- [0008]** Vorteilhafte Ausgestaltungen sind in den Unteransprüchen definiert.
- [0009]** Im Rahmen der vorliegenden Erfindung wird jetzt eine DNA-Sequenz offenbart, die für ein synthetisches Faserprotein, insbesondere ein synthetisches Spinnenseidenprotein kodiert, das eine mindestens 80%ige, vorzugsweise mindestens 84%ige, mehr bevorzugt mindestens 88%ige, besonders bevorzugt mindestens 90%ige und 92%ige, am meisten bevorzugt mindestens 94%ige Homologie zu Spidroin- und/oder Fibroin-Proteinen, insbesondere zu dem Spidroin 1-Protein, besonders bevorzugt zu dem Spidroin 1-Protein aus *Nephila clavipes* aufweist.
- [0010]** Homologie bedeutet im Rahmen dieser Erfindung Ähnlichkeit zwischen Aminosäuresequenzen aufgrund von identischen bzw. homologen Aminosäurebausteinen. Welche Aminosäuren als homolog anzusehen sind, ist dem Fachmann bekannt, z. B. (i) Isoleucin, Leucin und Valin untereinander, (ii) Asparagin und Glutamin, (iii) Asparaginsäure und Glutaminsäure.
- [0011]** Die erfindungsgemäße DNA-Sequenz ist aus Modulen aufgebaut, die eine Gruppe von aneinandergereihten Oligonukleotidsequenzen umfassen, wobei die Oligonukleotidsequenzen jeweils für repetitive Einheiten aus Spidroin- und/oder Fibroin-Proteinen kodieren.
- [0012]** Der Aufbau der erfindungsgemäßen DNA-Sequenz aus verschiedenen Modulen, welche wiederum aus unterschiedlichen, für Spidroine bzw. Fibroine typischen kurzen Aminosäure-Repeats konstruiert sind, wobei sich das Prinzip der Aneinanderreihung der entsprechenden Oligonukleotidsequenzen bzw. der Module an natürlichen Spidroin- und/oder Fibroin-Sequenzen orientiert, gewährleistet eine sehr hohe Homologie zu bisher bekannten natürlichen Spidroin- bzw. Fibroin-Sequenzen. Dadurch wird sichergestellt, daß die durch die erfindungsgemäße DNA-Sequenz kodierten Spinnenseidenproteine nach dem Verspinnen zu Fäden hervorragende mechanische Eigenschaften bezüglich ihrer Festigkeit und Elastizität aufweisen, die mit den mechanischen Eigenschaften von natürlichen Spinnenfäden vergleichbar sind.
- [0013]** Des weiteren ermöglicht der modulartige Aufbau der erfindungsgemäßen DNA-Sequenz eine einfache gentechnische Modifizierung der synthetischen Gene, so daß Multimeren von synthetischen Spinnenseidenproteinen mit beliebiger Größe je nach Wunsch hergestellt werden können. Ferner können die durch die erfindungsgemäße DNA-Sequenz kodierten Spinnenseidenproteine aufgrund des modulartigen Aufbaus mit anderen Faserproteinsequenzen fusioniert werden. Ein besonderer Vorteil der erfindungsgemäßen DNA-Sequenz ist, daß sie aufgrund ihres modulartigen Aufbaus auf einfache Weise mit für Reinigungselemente oder Löslichkeits-verändernde Peptide kodierenden Sequenzen fusioniert werden kann.
- [0014]** Bei einer besonders bevorzugten Ausführungsform der vorliegenden Erfindung weist das durch die erfindungsgemäße DNA-Sequenz kodierte Spinnenseidenprotein eine mindestens 84%ige, vorzugsweise mindestens 90%ige und besonders bevorzugt mindestens 94%ige Homologie zu dem Spidroin 1-Protein aus *Nephila clavipes* auf. Spidroin 1 aus *Nephila clavipes* ist wesentlich am Aufbau eines mechanisch besonders stabilen und elastischen Tragfadens beteiligt.

**[0015]** Aufgrund des modularen Aufbaus der erfindungsgemäßen DNA-Sequenz ist die Konstruktion von Genen, die sehr große Spinnenseidenproteine kodieren, ohne weiteres möglich, wobei die hohe Homologie zu Spidroin- und/oder Fibroin-Proteinen, insbesondere zu Spidroin 1, besonders bevorzugt zu Spidroin 1 von *Nephila clavipes* immer erhalten bleibt. Die so erzielbare Größenverteilung der durch die erfindungsgemäßen DNA-Sequenzen kodierten Proteine entspricht dem Spektrum von Spinnenseidenproteinen, das nach der Auflösung von natürlicher Spinnenseide beobachtet werden kann. Dieses identische Größenspektrum sowie die hohe Sequenzhomologie definieren die erfindungsgemäßen synthetischen Gene als Gene, die für Spinnenseidenproteine kodieren. Im Gegensatz zu natürlicher Spinnenseide, die aus einem Gemisch von Spinnenseidenproteinen besteht, werden durch die vorliegende Erfindung Spinnenseidenprotein-Gene bereitgestellt, die mit hoher Homologie eine Genklasse repräsentieren und eine einfache gentechnische Manipulation erlauben.

**[0016]** Die Module zum Aufbau der erfindungsgemäßen DNA-Sequenz umfassen eine Gruppe von aneinandergereihten Oligonukleotidsequenzen, die vorzugsweise ausgewählt sind aus der Gruppe, bestehend aus:

- |  |    |
|--|----|
| a) TATGAGCGCTCCCGGGCAGGGT;   |    |
| b) AGCTTTTAGGTACCAATATTAATCTGGCCGGCTCCACC;   | 15 |
| c) TATGGTCTGGGG;   |    |
| d) GGCCAGGGTGCTGGCCAA;   |    |
| e) GGTGCAGGAGCWGCWGCWGCWGCCTGCAGGTGGA;   |    |
| f) GCCGGCCAGATTAATATTGGTACCTAAA;   |    |
| g) CTGCCCCGGGAGCGCTCA;   | 20 |
| h) ACCACCATAACCTCC;  |    |
| i) AGCACCCCTGGCCCCCAG;   |    |
| j) TGCAGCWGCWGCWGCWGCCTCTGCACCTTGGCC;  |    |
| k) TATGAGATCTGGCCAAGGAGGT;   |    |
| l) TTGGCCAGATCTCA;   | 25 |
| m) AGTCAGGGTGCTGGTCGTGGAGGCCAA;  |    |
| n) TCCACGACCAGCACCCCTGACTCCCCAG;   |    |
| o) AGTCAGGGCGCTGGTCGTGGGGGACTGGGTGGCCAA;   |    |
| p) ACCCAGTCCCCCACGACCAGCGCCCTGACTCCCCAG;   |    |
| q) CTGGGAGGGCAGGGAGCGGGCCAA;   | 30 |
| r) CGCTCCCTGCCCCTCCAGACCTCC; und   |    |
| s) Sequenzen, die zu den Sequenzen a) bis r) eine mindestens 80%ige, vorzugsweise mindestens 90%ige, besonders bevorzugt mindestens 94%ige Sequenzidentität aufweisen. |    |

**[0017]** Die Module umfassen vorzugsweise mindestens vier Oligonuklotidsequenzen, die sich vorzugsweise unterscheiden, um die natürlichen Spinnenseidenproteine auf authentische Weise nachzugestalten. Die erfindungsgemäße DNA-Sequenz ist wiederum vorzugsweise aus mindestens vier der vorstehend beschriebenen Module aufgebaut.

**[0018]** Der Aufbau der erfindungsgemäßen DNA-Sequenz wird im folgenden beispielhaft dargestellt. Zunächst werden die in **Abb. 1** angegebenen Oligonukleotide bereitgestellt, die für Aminosäuresequenzen kodieren, die Spidroin-typischen kurzen Aminosäure-Repeats entsprechen. Diese Oligonukleotide werden durch gentechnische Verfahren miteinander kombiniert, wobei sich die Kombination an der natürlichen Spidroin-Sequenz richtet (siehe **Abb. 2**). Die so entstandenen Module A, B, C, D, E und F werden erneut miteinander kombiniert (siehe **Abb. 3**). Auf diese Weise werden erfindungsgemäße DNA-Sequenzen bereitgestellt, die auf Aminosäureebene eine mindestens 85%ige, vorzugsweise mindestens 90%ige und besonders bevorzugt mindestens 94%ige Homologie zu Spidroin-Proteinen zeigen.

[0019] Bei einer weiteren Ausführungsform umfasst die erfindungsgemäße DNA-Sequenz zusätzlich zu den vorstehend beschriebenen Modulen Nukleinsäuresequenzen, die für repetitive Einheiten aus Fibroin-Proteinen, vorzugsweise aus dem Fibroin-Protein des Seidenspinners kodieren.

[0020] Besonders bevorzugte erfindungsgemäße DNA-Sequenzen weisen die Sequenzen SEQ ID No. 19 bis 29 auf.

[0021] Erfindungsgemäß ist es ferner überraschenderweise erstmals gelungen, synthetische Spinnenseidenproteine in transgenen Pflanzen zu erzeugen. Auf diese Weise können synthetische Spinnenseidenproteine in großem Maßstab hergestellt werden. Um eine stabile Expression der erfindungsgemäßen DNA-Sequenz in Pflanzen zu gewährleisten, wird erfindungsgemäß ein rekombinantes Nukleinsäuremolekül bereitgestellt, das die vorstehend beschriebene erfindungsgemäße DNA-Sequenz sowie einen ubiquitär wirkenden Promotor, vorzugsweise den CaMV35S-Promotor umfaßt. Die Bereitstellung des erfindungsgemäßen rekombinanten Nukleinsäuremoleküls ermöglicht die Expression und Akkumulation von synthetischen Spidroin- bzw. Fibroinsequenzen in transgenen Pflanzen.

**[0022]** Um sicherzustellen, daß die erfindungsgemäße DNA-Sequenz in geeigneten Kompartimenten von transgenen Pflanzen exprimiert und akkumuliert wird, umfaßt das erfindungsgemäße Nukleinsäuremolekül zusätzlich zu der erfindungsgemäßen DNA-Sequenz und einem ubiquitär wirkenden Promotor vorzugsweise mindestens eine Nukleinsäuresequenz, die für ein pflanzliches Signalpeptid kodiert.

[0023] Bei einer bevorzugten Ausführungsform wird als Zielkompartiment für die Expression bzw. Akkumulation des synthetischen Spinnenseidenproteins das endoplasmatische Retikulum (ER) ausgewählt. Dieses Kompartiment ist für die stabile Akkumulation von Fremdproteinen in Pflanzen besonders geeignet. Um den Transport in das ER zu gewährleisten, umfaßt das erfindungsgemäße Nukleinsäuremolekül bevorzugt entsprechende Signalpeptide, besonders bevorzugt die LeB4Sp-Sequenz.

[0024] Die Retention im ER, falls gewünscht, wird erfindungsgemäß dadurch gewährleistet, daß das erfindungsgemäße Nukleinsäuremolekül zusätzlich eine Nukleinsäuresequenz umfaßt, die für ein ER-Retentionspeptid kodiert. Vorzugsweise wird die Retention in ER durch die C-terminal angefügte Aminosäuresequenz KDEL erreicht.

[0025] Ferner kann es vorteilhaft sein, die erfindungsgemäße DNA-Sequenz an der Plasmamembra, d. h. der Zellmem-

bran, zu plazieren. Deshalb umfaßt das erfindungsgemäße rekombinante Nukleinsäuremolekül bei einer alternativen Ausführungsform die erfindungsgemäße DNA-Sequenz, fusioniert an den N-Terminus einer Transmembrandomäne. Vorzugsweise ist diese Transmembrandomäne die Transmembrandomäne des PDGF-Rezeptors, die sogenannte HOOK-Sequenz (siehe Abb. 4).

- 5 **[0026]** Bei einer besonders bevorzugten Ausführungsform der vorliegenden Erfindung ist das erfindungsgemäße Nukleinsäuremolekül mit ELP's (elastin-like polypeptides) fusioniert. ELP's sind oligomere Repeats des Pentapeptids Val-Pro-Gly-Xaa-Gly (wobei Xaa jede Aminosäure außer Prolin und vorzugsweise Gly ist) und unterliegen einem reversiblen inversen Temperaturübergang. Sie sind in Wasser unterhalb der inversen Übergangstemperatur ( $T_i$ ) sehr gut löslich, unterliegend jedoch einem scharfen Phasenübergang im Bereich von 2°C bis 3°C, wenn die Temperatur über  $T_i$  erhöht wird, was zur Ausfällung und Aggregation des Polypeptids führt. D. E. Meyer und A. Chilkoti, Nat. Biotech. 1999, 17, 1112-1115, haben beschrieben, daß ELP-Fusionen mit rekombinanten Proteinen das Löslichkeitsverhalten dieser rekombinanten Proteine bei verschiedenen Temperaturen und Konzentrationen gezielt verändern. Bei der vorliegenden Erfindung wird dies zur Etablierung von im nachfolgenden detailliert beschriebenen Reinigungsstrategien für das durch die erfindungsgemäße DNA-Sequenz kodierte Spinnenseidenprotein genutzt. Vorzugsweise umfassen die durch die Nukleinsäuresequenz in dem erfindungsgemäßen Nukleinsäuremolekül kodierten ELP's von 10 bis 100 der vorstehend beschriebenen Pentamer-Einheiten (siehe Abb. 5).

- 10 **[0027]** Die Herstellung der vorstehend beschriebenen chimären Genkonstrukte bzw. rekombinanten Nukleinsäuremoleküle erfolgt mittels konventioneller Klonierungstechniken (siehe beispielsweise Sambrook et al. (1989, Molecular Cloning: A Laboratory Manual, 2. Auflage, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York)). Mittels dieser gängigen molekularbiologischen Techniken ist es möglich, gewünschte Konstrukte für die Transformation von Pflanzen vorzubereiten bzw. herzustellen. Die für die gentechnische Manipulation in prokaryontischen Zellen üblicherweise eingesetzten Klonierungs-, Mutagenisierungs-, Sequenzanalyse- und Restriktionsanalyse-Methoden sowie weitere biochemisch-molekularbiologische Methoden sind dem Fachmann wohlbekannt. So können nicht nur geeignete chimäre Genkonstrukte mit der jeweils gewünschten Fusion von Promotor, erfindungsgemäßer DNA-Sequenz, für ein pflanzliches Signalpeptid kodierender Sequenz, für ein ER-Retentionspeptid kodierender Sequenz, für eine Transmembrandomäne kodierender Sequenz und/oder für Reinigungselemente bzw. Löslichkeits-verändernde Peptide kodierenden Sequenzen hergestellt werden. Vielmehr kann der Fachmann mittels Routinetechniken, falls erwünscht, verschiedenartige Mutationen oder Deletionen in die jeweiligen Gene einführen.

- 25 **[0028]** Die Erfindung betrifft weiterhin Vektoren und Mikroorganismen, die erfindungsgemäße Nukleinsäuremoleküle enthalten und deren Verwendung die Herstellung von Pflanzenzellen bzw. Pflanzen ermöglicht, die Spinnenseidenproteine produzieren. Dabei handelt es sich bei den Vektoren insbesondere um Plasmide, Cosmide, Viren, Bakteriophagen und andere in der Gentechnik gängige Vektoren. Bei den Mikroorganismen handelt es sich in erster Linie um Bakterien, Viren, Pilze, Hefen und Algen.

- 30 **[0029]** Da die erfindungsgemäßen DNA-Sequenzen aufgrund ihres repetitiven Charakters kaum unikale Restriktionsorte aufweisen, wurden die erfindungsgemäßen Vektoren bzw. die das synthetische Spinnenseidenprotein kodierenden Gene durch verschiedene Strategien entsprechend angepaßt (siehe Abb. 6 bis 8). Bei der Amplifizierung der erfindungsgemäßen DNA-Sequenzen durch PCR werden aufgrund des extrem repetitiven Charakters der erfindungsgemäßen DNA-Sequenzen vorzugsweise zunächst Oligonukleotide anligiert, welche dann als Matrizen für die nachfolgenden PCR-Reaktionen dienen (siehe Abb. 7).

- 40 **[0030]** Des weiteren wird bei der vorliegenden Erfindung ein rekombinantes Spinnenseidenprotein bereitgestellt, das durch die erfindungsgemäße DNA-Sequenz kodiert wird. Dieses erfindungsgemäße synthetische Spinnenseidenprotein, vorzugsweise mit einem Molekulargewicht im Bereich von 10 bis 160 kDa, weist eine mindestens 85%ige, vorzugsweise mindestens 90%ige und besonders bevorzugt mindestens 94%ige Homologie zu Spidroin- und/oder Fibroin-Proteinen auf. Durch diese hohe Homologie mit den natürlichen Faserproteinen der Spinne und des Seidenspinners wird gewährleistet, daß die herausragenden mechanischen Eigenschaften der natürlichen Spinnenfäden erreicht werden, wenn die erfindungsgemäßen Proteine zu Fäden gesponnen werden.

- 45 **[0031]** Ferner weisen die erfindungsgemäßen Proteine überraschenderweise neuartige physikochemische Eigenschaften auf. So bleibt die Löslichkeit dieser erfindungsgemäßen synthetischen Faserproteine in wäßrigen Lösungen auch nach längerem Kochen außerordentlich gut erhalten. Gemeinsam mit der ebenfalls auftretenden Löslichkeit in organischen Lösungen und dem Fällungsverhalten bei hohen Salzkonzentrationen können diese neuen Eigenschaften der erfindungsgemäßen synthetischen Spinnenseidenproteine somit für die Entwicklung technisch durchführbarer Extraktions- und Reinigungsverfahren genutzt werden. Diese Eigenschaften werden noch verstärkt, wenn die erfindungsgemäßen synthetischen Spinnenseidenproteine gezielt in bestimmten Kompartimenten, insbesondere im ER von transgenen Pflanzen akkumuliert werden.

- 55 **[0032]** Beispiele für Aminosäuresequenzen der erfindungsgemäßen rekombinanten synthetischen Spinnenseidenproteine weisen die Sequenzen SEQ ID No. 30 bis 40 auf. Die erfindungsgemäßen Spinnenseidenproteine können alternativ auch nach chemischen, dem Fachmann bekannten Methoden synthetisiert werden, eine rekombinante Herstellung ist jedoch bevorzugt.

- 60 **[0033]** Ferner betrifft die Erfindung ein Verfahren zur Herstellung von Spinnenseidenprotein-produzierenden Pflanzen bzw. Pflanzenzellen, umfassend die folgenden Schritte:

- a) Herstellung eines wie vorstehend beschriebenen erfindungsgemäßen rekombinanten Nukleinsäuremoleküls;
- b) Übertragung des Nukleinsäuremoleküls aus a) auf pflanzliche Zellen; und
- c) gegebenenfalls die Regeneration fertiler Pflanzen aus den transformierten Pflanzenzellen.

- 65 **[0034]** Des weiteren betrifft die Erfindung Pflanzenzellen, die die erfindungsgemäßen Nukleinsäuremoleküle bzw. den erfindungsgemäßen Vektor enthalten. Die Erfindung betrifft ferner Ernteprodukte und Vermehrungsmaterial transgener Pflanzen sowie die transgenen Pflanzen selbst, die ein erfindungsgemäßes Nukleinsäuremolekül enthalten.

- [0035] Zur Vorbereitung der Einführung fremder Gene in höhere Pflanzen bzw. deren Zellen stehen eine große Anzahl von Klonierungsvektoren zur Verfügung, die ein Replikationssignal für *E. coli* und ein Markergen zur Selektion transformierter Bakterienzellen enthalten. Beispiele für derartige Vektoren sind pBR322, pUC-Serien, M13mp-Serien, pA-CYC184 usw. Die gewünschte Sequenz kann an einer passenden Restriktionsschnittstelle in den Vektor eingeführt werden. Das erhaltene Plasmid wird dann für die Transformation von *E. coli*-Zellen verwendet. Transformierte *E. coli*-Zellen werden in einem geeigneten Medium gezüchtet und anschließend geerntet und lysiert, und das Plasmid wird wiedergewonnen. Als Analysenmethode zur Charakterisierung der gewonnenen Plasmid-DNA werden im allgemeinen Restriktionsanalysen, Gelelektrophoresen und weitere biochemisch-molekularbiologische Methoden eingesetzt. Nach jeder Manipulation kann die Plasmid-DNA gespalten und gewonnene DNA-Fragmente mit anderen DNA-Sequenzen verknüpft werden. 5
- [0036] Für die Einführung von DNA in eine pflanzliche Wirtszelle stehen eine Vielzahl von Techniken zur Verfügung, wobei der Fachmann die jeweils geeignete Methode ohne Schwierigkeiten ermitteln kann. Diese Techniken umfassen die Transformation pflanzlicher Zellen mit T-DNA unter Verwendung von *Agrobacterium tumefaciens* oder *Agrobacterium rhizogenes* als Transformationsmedium, die Fusion von Protoplasten, die Injektion, die Elektroporation, den direkten Gentransfer isolierter DNA in Protoplasten, die Einbringung von DNA mittels der biolistischen Methode sowie weitere Möglichkeiten die bereits seit mehreren Jahren gut etabliert sind und zum üblichen Repertoire des Fachmanns in der pflanzlichen Molekularbiologie bzw. Pflanzenbiotechnologie gehören. 10
- [0037] Bei der Injektion und Elektroporation von DNA in Pflanzenzellen werden per se keine speziellen Anforderungen an die verwendeten Plasmide gestellt. Ähnliches gilt für den direkten Gentransfer. Es können einfache Plasmide, wie z. B. pUC-Derivate, verwendet werden. Sollen aber aus derartig transformierten Zellen ganze Pflanzen regeneriert werden, ist die Anwesenheit eines selektierbaren Markergens empfehlenswert. Dem Fachmann sind die gängigen Selektionsmarker bekannt, und es stellt für ihn kein Problem dar, einen geeigneten Marker auszuwählen. 20
- [0038] Je nach Einführungsmethode gewünschter Gene in die Pflanzenzelle können weitere DNA-Sequenzen erforderlich sein. Werden z. B. für die Transformation der Pflanzenzelle das Ti- oder Ri-Plasmid verwendet, so muss mindestens die rechte Begrenzung, häufig jedoch die rechte und linke Begrenzung der im Ti- bzw. Ri-Plasmid enthaltenen T-DNA als Flankenbereich mit den einzuführenden Genen verbunden werden. Werden für die Transformation Agrobakterien verwendet, muss die einzuführende DNA in spezielle Plasmide kloniert werden, und zwar entweder in einen intermediären oder in einen binären Vektor. Die intermediären Vektoren können aufgrund von Sequenzen, die homolog zu Sequenzen in der T-DNA sind, durch homologe Rekombination in das Ti- oder Ri-Plasmid der Agrobakterien integriert werden. Dieses enthält außerdem die für den Transfer der T-DNA notwendige vir-Region. Intermediäre Vektoren können nicht in Agrobakterien replizieren. Mittels eines Helferplasmids kann der intermediäre Vektor auf *Agrobacterium tumefaciens* übertragen werden (Konjugation). Binäre Vektoren können sowohl in *E. coli* als auch in Agrobakterien replizieren. Sie enthalten ein Selektionsmarkergen und einen Linker oder Polylinker, welche von der rechten und linken T-DNA-Grenzregion eingerahmt werden. Sie können direkt in die Agrobakterien transformiert werden. Das als Wirtszelle dienende Agrobakterium soll ein Plasmid, das eine vir-Region trägt, enthalten. Die vir-Region ist für den Transfer der T-DNA in die Pflanzenzelle notwendig. Zusätzliche T-DNA kann vorhanden sein. Das derartig transformierte Agrobakterium wird zur Transformation von Pflanzenzellen verwendet. Die Verwendung von T-DNA für die Transformation von Pflanzenzellen ist intensiv untersucht und ausreichend in allseits bekannten Übersichtsartikeln und Handbüchern zur Pflanzen- 30 transformation beschrieben worden. Für den Transfer der DNA in die Pflanzenzelle können Pflanzen-Explantate zweckmäßigerweise mit *Agrobacterium tumefaciens* oder *Agrobacterium rhizogenes* kultiviert werden. Aus dem infizierten Pflanzenmaterial (z. B. Blattstücke, Stengelsegmente, Wurzeln, aber auch Protoplasten oder Suspensions-kultivierte Pflanzenzellen) können dann in einem geeigneten Medium, welches Antibiotika oder Biozide zur Selektion transformierter Zellen enthalten kann, wieder ganze Pflanzen regeneriert werden. 35
- [0039] Ist die eingeführte DNA einmal im Genom der Pflanzenzelle integriert, so ist sie dort in der Regel stabil und bleibt auch in den Nachkommen der ursprünglich transformierten Zelle erhalten. Sie enthält normalerweise einen Selektionsmarker, der den transformierten Pflanzenzellen Resistenz gegenüber einem Biozid oder einem Antibiotikum wie Kanamycin, G 418, Bleomycin, Hygromycin, Methotrexat, Glyphosat, Streptomycin, Sulfonylharnstoff, Gentamycin oder Phosphinotricin u. a. vermittelt. Der individuell gewählte Marker sollte daher die Selektion transformierter Zellen gegenüber Zellen, denen die eingeführte DNA fehlt, gestatten. Hierfür sind auch alternative Marker geeignet, wie nutritive Marker, Screeningmarker (wie GFP, green fluorescent protein). Selbstverständlich kann auch vollkommen auf Selektionsmarker verzichtet werden, was allerdings mit einem ziemlich hohen Screeningbedarf einhergeht. Falls markerfreie transgenen Pflanzen erwünscht sind, stehen dem Fachmann auch Strategien zur Verfügung, die eine nachträgliche Entfernung des Markergens erlauben, z. B. Cotransformation, Sequenz-spezifische Rekombinasen. 40
- [0040] Die Regeneration der transgenen Pflanzen aus transgenen Pflanzenzellen erfolgt nach üblichen Regenerationsmethoden unter Verwendung bekannter Nährmedien. Die so erhaltenen Pflanzen können dann mittels üblicher Verfahren, einschließlich molekularbiologischer Methoden, wie PCR, Blot-Analysen, auf Anwesenheit der eingeführten Nukleinsäure, die ein synthetisches Spinnenseidenprotein kodiert, untersucht werden. 45
- [0041] Bei der transgenen Pflanze bzw. der transgenen Pflanzenzelle kann es sich um jede beliebige monokotyle oder dikotyle Pflanze bzw. Pflanzenzelle handeln. Vorzugsweise handelt es sich um Nutzpflanzen bzw. Zellen von Nutzpflanzen. Besonders bevorzugt handelt es sich um transgene Pflanzen ausgewählt aus der Gruppe, bestehend aus der Tabakpflanze (*Nicotiana tabacum*) und der Kartoffelpflanze (*Solanum tuberosum*). 50
- [0042] Die Expression des erfindungsgemäßen synthetischen Spinnenseidenproteins in den erfindungsgemäßen Pflanzen bzw. in den erfindungsgemäßen Pflanzenzellen kann mit Hilfe herkömmlicher molekularbiologischer und biochemischer Methoden nachgewiesen und verfolgt werden. Dem Fachmann sind diese Techniken bekannt und er ist problemlos in der Lage, eine geeignete Nachweismethode zu wählen, beispielsweise eine Northern-Blot-Analyse oder eine Southern-Blot-Analyse. 55
- [0043] Ein Beispiel für die Herstellung von transgenen Spinnenseidenprotein-produzierenden Pflanzen ist in Abb. 9 angegeben. Die durch PCR amplifizierte Sequenzen können möglicherweise frameshift-Mutationen enthalten. Deshalb 60

müssen die erfindungsgemäßen Sequenzen vor der Erzeugung transgener Pflanzen überprüft werden. Dies kann durch Sequenzanalyse jeweils von den flankierenden Vektorsequenzen aus erfolgen. Längere Konstrukte über 1 kB können auf diese Weise nicht geprüft werden, da aufgrund der repetitiven Eigenschaften der erfindungsgemäßen DNA-Sequenzen interne Sequenzierungsprimer keine auswertbaren sicheren Sequenzen liefern. Aus diesem Grund wurden amplifizierte

5 Spidroinsequenzen vorzugsweise in den bakteriellen Expressionsvektor pet23a (Novagen, Madison, USA) kloniert. Durch immunchemischen Nachweis der Expression können dann frameshift-Mutationen ausgeschlossen werden.

[0044] Die erfindungsgemäßen Nukleinsäuremoleküle bzw. Expressionskassetten werden erfindungsgemäß üblicherweise als HindIII-Fragmente in Shuttle-Vektoren wie pBIN, pCB301 und/oder pGSGLOC1 kloniert. Diese Shuttle-Vektoren werden vorzugsweise in *Agrobacterium tumefaciens* transformiert. Die Transformation von *Agrobacterium tumefaciens* wird üblicherweise durch Southern-Blot-Analyse und/oder PCR-Screening überprüft.

10 [0045] Die Erfindung betrifft ebenfalls Vermehrungsmaterial und Ernteprodukte der erfindungsgemäßen Pflanzen, beispielsweise Früchte, Samen, Knollen, Wurzelstöcke, Sämlinge, Stecklinge usw.

[0046] Ferner betrifft die Erfindung ein Verfahren zur Gewinnung von pflanzlichem Spinnenseidenprotein, umfassend die folgenden Schritte:

- 15 a) die Übertragung eines erfindungsgemäßen rekombinanten Nukleinsäuremoleküls oder Vektors, der eine DNA-Sequenz erhält, die für ein synthetisches Spinnenseidenprotein kodiert, auf Pflanzenzellen;
- b) gegebenenfalls die Regeneration von Pflanzen aus den transformierten Pflanzenzellen;
- 20 c) die Verarbeitung der Pflanzenzellen aus a) bzw. der Pflanzen aus b) zur Gewinnung von pflanzlichem Spinnenseidenprotein.

[0047] Bei einem weiteren wesentlichen Aspekt der vorliegenden Erfindung werden Verfahren zur Gewinnung von rekombinant hergestellten Spinnenseidenproteinen bereitgestellt, die die Übertragung eines erfindungsgemäßen rekombinanten Nukleinsäuremoleküls oder Vektors, der eine DNA-Sequenz enthält, die für ein synthetisches Spinnenseidenprotein kodiert, auf beliebige Zellen, d. h. neben Pflanzenzellen beispielsweise auch bakterielle oder tierische Zellen, umfassen. Wesentliches Merkmal bei diesen erfindungsgemäßen Verfahren ist dabei der Schritt der Reinigung der rekombinant hergestellten Spinnenseidenproteine, bei dem u. a. deren besondere Eigenschaften hinsichtlich der Löslichkeit bei Erwärmung und/oder Säurezugabe ausgenutzt werden.

[0048] So erfolgt bei einer Ausführungsform des erfindungsgemäßen Verfahrens die Reinigung des rekombinant hergestellten Spinnenseidenproteins durch Hitzebehandlung des Zellextrakts, z. B. eines Pflanzensamen-Extrakts, und anschließende Abtrennung der denaturierten zelleigenen, z. B. der pflanzeigenen Proteine beispielsweise durch Zentrifugation. Dabei wird die Eigenschaft der rekombinant hergestellten Spinnenseidenproteine ausgenutzt, daß sie beim Erhitzen von wäßrigen Lösungen bis zum Siedepunkt löslich bleiben. Dagegen bleiben beispielsweise synthetische Faserproteine der Spinne und des Seidenspinners nach Expression in *Pichia pastoris* beim Erhitzen nur bis zu einer Temperatur

35 von 63°C und dann nur für 10 Minuten in Lösung.  
[0049] Bei einer weiteren Ausführungsform des erfindungsgemäßen Verfahrens zur Gewinnung von rekombinant hergestellten Spinnenseidenproteinen erfolgt die Reinigung durch Einstellung eines sauren pH mittels Zugabe von Säure, vorzugsweise Salzsäure zu dem Zellextrakt, beispielsweise zu dem Pflanzenextrakt. Der saure pH, insbesondere ein pH im Bereich von 1,0 bis 4,0, besonders bevorzugt im Bereich von 2,5 bis 3,5, am meisten bevorzugt ein pH von 3,0, wird dabei vorzugsweise für eine Dauer von mehreren Minuten, besonders bevorzugt etwa 30 Minuten, bei einer Temperatur unterhalb Raumtemperatur, vorzugsweise etwa 4°C, beibehalten. Wiederum wird eine nicht zu erwartende Eigenschaft der durch das erfindungsgemäße Verfahren gewonnenen Proteine ausgenutzt, nämlich daß sie beim Ansäuern, insbesondere bis zu einem pH von 3,0 bei 4°C in Lösung bleiben. Die zelleigenen, beispielsweise pflanzeigenen Proteine fallen dagegen aus und werden insbesondere durch Zentrifugation abgetrennt.

45 [0050] Die vorstehend beschriebenen Löslichkeitseigenschaften der nach dem erfindungsgemäßen Verfahren rekombinant hergestellten Spinnenseidenproteine sind sehr überraschend und waren in dieser Form nicht vorhersehbar und ermöglichen eine effiziente, schnelle und kostengünstige Reinigung bei deren Extraktion aus Zellen, insbesondere Pflanzenzellen.

[0051] Bei einer weiteren Ausführungsform des erfindungsgemäßen Verfahrens wird ein Nukleinsäuremolekül auf die Zellen übertragen, das zusätzlich eine Nukleinsäuresequenz umfaßt, die für ELP's kodiert. In diesem Fall erfolgt die Reinigung des rekombinant hergestellten Spinnenseidenproteins auf die folgende Weise: In einem ersten Schritt wird das Spinnenseiden-ELP-Fusionsprotein durch Hitzebehandlung des Rohextrakts angereichert. Überraschenderweise behalten die Fusionsproteine dabei die außerordentliche Löslichkeit der Spinnenseidenproteine bei hohen Temperaturen bei. Ein Großteil der zelleigenen Proteine fällt bei dieser Temperaturerhöhung aus. Im nächsten Schritt werden die Spinnenseiden-ELP-Fusionsproteine durch weitere Temperaturerhöhung, vorzugsweise auf eine Temperatur von mindestens 60°C, ausgefällt. Vorzugsweise erfolgt die Ausfällung bei einer geeigneten Salz-Konzentration, beispielsweise einer NaCl-Konzentration von mindestens 0,5 M, vorzugsweise im Bereich von 1 M bis 2 M. Schließlich wird das ELP-Fragment, vorzugsweise durch Verdauung mit CNBr, abgespalten.

60 [0052] Durch das vorstehend beschriebene erfindungsgemäße Verfahren zur Gewinnung von rekombinant hergestelltem Spinnenseidenprotein können die Proteine in Pflanzen zu hohen Konzentrationen, vorzugsweise bis zu einer Expressionshöhe von etwa 4% des gesamten löslichen Proteins angereichert werden. Damit werden erstmals Verfahren bereitgestellt, die zur technisch realisierbaren Anreicherung von rekombinantem Spinnenseidenprotein verwendet werden können.

65 [0053] Die erfindungsgemäßen Spinnenseidenproteine können bei einem weiteren Aspekt der vorliegenden Erfindung zur Herstellung von synthetischen Fäden sowie von Folien und Membranen verwendet werden. Insbesondere sind derartige Produkte für medizinische Anwendungen, insbesondere zum Vernähen von Wunden und/oder als Gerüste bzw. als Abdeckung für künstliche Organe, geeignet. Ferner können die aus den erfindungsgemäßen Spinnenseidenproteine hergestellten Folien und Membrane u. a. als Anheftungsflächen für kultivierte Zellen sowie zur Filterung verwendet werden.

den.

[0054] Die vorliegende Erfindung wird in den nachfolgenden Beispielen, die der Veranschaulichung der Erfindung dienen und in keiner Weise als einschränkend zu verstehen sind, erläutert.

## Beispiele

5

### Beispiel 1

[0055] Expression und stabile Akkumulation von synthetischen Faserproteinen der Spinne und des Seidenspinners im endoplasmatischen Retikulum von Blättern bzw. Knollen transgener Tabak- und Kartoffelpflanzen.

10

[0056] In den Abb. 10a und b sind Aminosäuresequenzen von synthetischen Spinnenseidenproteinen mit einer hohen Homologie zu dem Spidroin 1-Protein aus *Nephila clavipes* dargestellt, wobei der C-terminale und nicht repetitive konstante Bereich nicht abgebildet sind. Diese synthetischen Spinnenseidenproteine bestehen aus Modulen, die wiederum aneinandergereihte Oligonukleotidsequenzen umfassen. Durch Kombination mehrerer Module wurden die verschiedenen synthetischen Gene assembliert, wobei auch Mischformen mit Fibroin 1-nachempfundenen Sequenzen erzeugt wurden.

15

[0057] In der nachfolgenden Tabelle 1 sind verschiedene Pflanzenexpressionskassetten aufgelistet, die für verschiedene erfindungsgemäße synthetische Faserproteine mit den Sequenzen SEQ ID No. 30 bis 40 kodieren.

Tabelle 1

20

Pflanzenexpressions-Kassette	Anzahl der Aminosäuren (mit Leadersequenz)	Berechnetes Molekulargewicht (mit Leadersequenz)	Homologie
SB1 (SEQ ID No. 19)	Nr. 1 – 149 AS	11 kDa	Spidroin 1
SD1 (SEQ ID No. 21)	Nr. 2 – 182 AS	13 kDa	Spidroin 1
SA1 (SEQ ID No. 26)	Nr. 3 – 215 AS	16 kDa	Spidroin 1
SE1 (SEQ ID No. 20)	Nr. 4 – 275 AS	20 kDa	Spidroin 1
SF1 (SEQ ID No. 29)	Nr. 5 – 317 AS	24 kDa	Spidroin 1
SM12 (SEQ ID No. 28)	Nr. 6 – 410 AS	31 kDa	Spidroin 1
SO1 (SEQ ID No. 27)	Nr. 7 – 676 AS	52 kDa	Spidroin 1
SO1SM12 (SEQ ID No. 23)	Nr. 8 – 1035 AS	82 kDa	Spidroin 1
SO1SO1 (SEQ ID No. 22)	Nr. 9 – 1301 AS	102 kDa	Spidroin 1
SO1SO1SO1 (SEQ ID No. 24)	Nr. 10 – 1926 AS	151 kDa	Spidroin 1
FA2 (SEQ ID No. 25)	Nr. 11 – 264 AS	20 kDa	Spidroin 1 und Fibroin

25

30

35

40

45

50

[0058] Durch eine N-terminale Signalpeptidsequenz und eine C-terminale ER-Retentionssequenz (KDEL) wurde der zielgerichtete Transport und die Akkumulation der erfindungsgemäßen Sequenzen im endoplasmatischen Retikulum von Zellen transgener Pflanzen erreicht. Eine Nachweissequenz in Form eines c-myc-Tags am C-terminalen Ende der transgenen synthetischen Faserproteine erlaubt den Nachweis der transgenen Produkte in Pflanzenextrakten.

55

[0059] Die Kassetten SO1 und FA2 sind beispielhaft in den Abb. 10a und 10b im Detail dargestellt. Nach demselben Aufbauprinzip wurden die Pflanzenexpressionskassetten SB1, SD1, SA1, SE1, SF1, SM12, SO1SM12, SO1SO1 und SO1SO1SO1 erstellt. Durch Variation der Grundmodulwiederholungen entstehen synthetische Faserproteine verschiedener Aminosäureanzahl und entsprechend unterschiedlichen Molekulargewichts (siehe Tabelle 1).

60

[0060] Die Abb. 2 beschreibt schematisch den Weg zur Einstellung der oben genannten Konstrukte. Zur direkten Klonierung der erfindungsgemäßen synthetischen Faserproteingene wurden die Schnittstellen *Sma*I und *Nae*I eingeführt. Dazu wurde ein PCR-Produkt, welches die entsprechenden Schnittstellen enthielt, mit der Primerkombination 5'-pRTRA-*Sma*I und 3'-pRTRA-*Not*I in das Plasmid pRTRA ScFv *Sma*Δ/*Bam*HIΔ über *Bam*HI und *Not*I kloniert. Synthetische Faserproteingene wurden aus den Faserproteingenderivaten der Plasmide 9905 oder 9609 in den Vektor pRTRA.7/3-Platzhalter kloniert. Durch die Wahl der Restriktionsendonukleaseerkennungssequenzen am 5'- und 3'-Ende der synthetischen Faserproteingene (*Sma*I und *Nae*I) sind diese frei miteinander kombinierbar, und größere Faserproteingene können erfindungsgemäß in einem Klonierungsschritt assembliert werden.

65



- [0061] Auf diese Weise wurden transgene synthetische Spinnenseidenproteine zu hohen Konzentrationen im endoplasmatischen Retikulum transgener Tabak- und Kartoffelpflanzen akkumuliert (siehe Abb. 12a und 12b). In der folgenden Tabelle 2 ist die maximale Akkumulationshöhe von erfindungsgemäßen synthetischen Spinnenseidenproteinen im ER von Blättern transgener Tabak- und Kartoffelpflanzen dargestellt. Die Abschätzung der Anreicherung der transgenen synthetischen Faserproteine erfolgte über einen Vergleich mit transgenen rekombinanten Antikörpern, die auf identische Weise mit dem gleichen Tag versehen wurden. Damit wird erstmals eine Akkumulation von Spinnenseidenproteinen in Pflanzen am Beispiel von Kartoffeln und Tabak beschrieben.

Tabelle 2

	Faserprotein			
	SD1	SM12	SO1	FA2
Tabak				
Akkumulationsmenge in Prozent Gesamtprotein	~ 0,5 %	~ 0,5 %	~ 0,5 %	~ 0,5 %
Kartoffel				
Akkumulationsmenge in Prozent Gesamtprotein	~ 0,5 %	~ 0,5 %	~ 0,5 %	~ 0,5 %

- [0062] Eine definierte Menge des faserproteinhaltigen Gesamtproteinextrakts (40 µg) und eine definierte Menge eines Referenzproteins mit c-myc-Immunotag (50 ng ScFv) wurden mittels SDS-Gelelektrophorese aufgetrennt, und synthetische Faserproteine und Referenzproteine wurden im Western Blot durch einen Anti-c-myc-Antikörper nachgewiesen (siehe Abb. 12 und 13). Die prozentualen Angaben resultieren aus dem Vergleich zwischen der Bandenintensität der Referenzproteine und der Bandenintensität der synthetischen erfindungsgemäßen Spinnenseidenproteine und stellen geschätzte Werte dar. Größenunterschiede der synthetischen Faserproteine und des Referenzproteins wurden berücksichtigt. Mögliche Unterschiede in der Markierungseffizienz können nahezu ausgeschlossen werden.

- [0063] In Abb. 13 ist die Hitzestabilität von verschiedenen erfindungsgemäßen synthetischen Spinnenseidenproteinen in pflanzlichen Extrakten dargestellt. Überraschenderweise bleiben die erfindungsgemäßen Spinnenseidenproteine auch bei einer längeren Hitzebehandlung von 3 h in Lösung (Vergleich der Referenzprobe R zu den Proben H-60 min, H-120 min und H-180 min). Die übrigen pflanzlichen Proteine werden zu mehr als 90% denaturiert und können durch Zentrifugation einfach abgetrennt werden (Abb. 13a: Vergleich der Probe R zu H-60 min). Diese ungewöhnlichen Eigenschaften der erfindungsgemäßen synthetischen Spinnenseidenproteine, die unter anderem durch ihre Aminosäuresequenz und ihre Faltung im pflanzlichen ER bedingt sind, ermöglichen die Entwicklung von großtechnisch realisierbaren und kostengünstigen Reinigungsstrategien.

- [0064] In Abb. 14 wird die Löslichkeit von synthetischen Faserproteinen aus transgenen Pflanzen dargestellt. Im Gegensatz zu den im Stand der Technik beschriebenen bakteriell exprimierten synthetischen Faserproteinen weisen die erfindungsgemäßen Spinnenseidenproteine eine überraschend gute Löslichkeit in wässrigen Puffern auf (R1, R2 = Tris-Puffer; T1, T2 = Phosphatpuffer). Auch diese Eigenschaften beruhen unter anderem auf der Aminosäuresequenz und insbesondere auf der Faltung im endoplasmatischen Retikulum pflanzlicher Zellen.

## Beispiel 2

- [0065] Expression und stabile Akkumulation von synthetischen Spinnenseidenproteinen im Plasmalemma von Blättern transgener Tabak- und Kartoffelpflanzen.

- [0066] In diesem Beispiel wird die membranassoziierte Akkumulation von erfindungsgemäßen Spinnenseidenproteinen in transgenen Tabak- und Kartoffelpflanzen beschrieben. Dabei wurden ausgehend von den in Beispiel 1 beschriebenen Konstrukten Fusionsgene hergestellt, die für ein Spinnenseidenprotein und für eine Membrandomäne kodieren. Das allgemeine Schema dieser Konstruktionen ist in Abb. 15 dargestellt. Dabei wurde aus dem Plasmid pRT-HOOK ein NotI-Fragment isoliert, welches sowohl für die HOOK-Domäne als auch für einen c-myc-Immunotag kodiert, welches dann in Spinnenseidenproteingen-tragende Derivate des Vektors pRTA.7/3 kloniert wurde. Durch die Wahl der Restriktionsendonukleaseerkennungssequenzen am 5'- und 3'-Ende der synthetischen Spinnenseidenproteingene (SmaI und NaeI) sind diese wiederum miteinander kombinierbar, wodurch größere Faserproteingene in einem Klonierungsschritt assimiliert werden können.

- [0067] Abb. 16 zeigt die Expression der vorstehend beschriebenen Gene in transgenen Tabak- und Kartoffelpflanzen. Wie aus dem Vergleich der Proben 1, 2 und 3 in dieser Abbildung ersichtlich ist, sind diese transgenen Spinnenseidenproteine im Gegensatz zu den im Beispiel 1 beschriebenen erfindungsgemäßen Proteinen in der wässrigen Phase nicht löslich. Auch diese Eigenschaft kann für die Entwicklung von Reinigungsstrategien ausgenutzt werden.

## Beispiel 3

- [0068] Gezielte Veränderung der Löslichkeit von Spinnenseidenproteinen durch Fusion mit elastin-like peptides.

- [0069] In einem ersten Schritt wurde gezeigt, daß Fusionen mit elastin-like peptides auch bei bakteriell exprimierten Spinnenseidenproteinen zu einer gezielten Veränderung des Löslichkeitsverhaltens in Abhängigkeit von Temperatur und Konzentration führen.

- [0070] Eine entsprechende Expressionskassette ist in **Abb. 5** dargestellt. Beispiele für ELP mit 10, 20, 30, 40, 60, 70 und 100 Pentamerseinheiten sind in den Sequenzen SEQ ID No. 41 bis 47 angegeben. Beispiele für DNA-Sequenzen und Aminosäuresequenzen in Form des Konstrukts SM12-70xELP als Pflanzenexpressionskassette bzw. als Expressionskassette für *E. coli* sind in den Sequenzen SEQ ID No. 48–51 bzw. in den **Abb. 19 bis 22** angegeben.
- [0071] In **Abb. 17** wird die gelelektrophoretische Analyse eines solchen Reinigungsverfahrens dargestellt. Durch Hitzebehandlung des Rohextrakts wurde das Spinnenseiden-ELP-Fusionsprotein angereichert. Überraschenderweise behielten die Fusionsproteine die außerordentliche Löslichkeit der Spinnenseidenproteine bei hohen Temperaturen bei. Ein Großteil der *E. coli*-Proteine wurde bei diesen Temperaturen ausgefällt.
- [0072] Nach starker Konzentration des angereicherten Spinnenseidenproteinextrakts wurde das Extrakt einer Temperatur von 60°C ausgesetzt, woraufhin das ELP-Spinnenseidenprotein ausfiel und pelletiert wurde. Das Pellet wurde bei Raumtemperatur in Wasser gelöst, und unlösliche Bestandteile wurden durch Pelletieren entfernt.
- [0073] Anschließend wurde die Spinnenseidenproteinfraktion lyophilisiert und durch Cyanbromidspaltung verdaut. Die Cyanbromidspaltung wurde durch den Methionin-Rest zwischen dem Spinnenseidenprotein und dem ELP-Peptid ermöglicht.
- [0074] Anschließend wurde erneut lyophilisiert und in wäßrigem Puffer gelöst. Dann erfolgte eine starke Konzentration, wobei das abgespaltene ELP-Fragment (ELP(T-R); siehe **Abb. 2**) ausfiel und durch Pelletieren entfernt wurde. Das Spinnenseidenprotein blieb dabei in Lösung (SM12(T-R); siehe **Abb. 17**). Die Löslichkeit blieb für einen längeren Zeitraum erhalten, bei SM12 bei 4°C für 24 H. Die Identität des auf diese Weise gereinigten Spinnenseidenproteins wurde durch Peptidsequenzierung des N-terminalen Endes gezeigt.
- [0075] In einem zweiten Schritt wurden Spinnenseidenproteine als ELP-Fusionen im endoplasmatischen Retikulum von transgenen Tabakpflanzen akkumuliert. Der prinzipielle Aufbau dieser Expressionskassetten ist ebenfalls in **Abb. 5** dargestellt. Diese Fusionsproteine mit Molekulargewichten von 35.000 Dalton bis 100.000 Dalton wurden sämtlich in Pflanzen zu hohen Konzentrationen mit einer Expressionshöhe von etwa 4% des gesamt löslichen Proteins angereichert.
- Abbildungen 25
- Abb. 1**
- [0076] Oligonukleotid-Sequenzen, die für Spidroin-typische kurze Aminosäurerepeats kodieren.
- 30
- Abb. 2**
- [0077] Aneinanderreihung von Oligonukleotid-Sequenzen zum Aufbau von Modulen der erfindungsgemäßen DNA-Sequenzen.
- 35
- Abb. 3**
- [0078] Aufbau der erfindungsgemäßen DNA-Sequenzen aus Modulen.
- 40
- Abb. 4**
- [0079] Klonierung des Gens der Transmembrandomäne von HOOK mit NotI aus (pRT-HOOK) in (pRTA.73 syn.spidroin).
- 45
- Abb. 5**
- [0080] Schematische Darstellung der Spidroin-ELP-Expressionskassetten. xELP-Einheiten: 10, 20, 30, 40; 60, 70 oder 100 Pentamere (Val-Pro-Gly-Val-Gly). Das Methionin zwischen dem Spinnenseidenprotein und dem ELP-Peptid ermöglicht die Cyanbromidspaltung.
- 50
- Abb. 6**
- [0081] Veränderung einer Base in der Erkennungssequenz von BamHI (Position 1332) durch gezielte Mutagenese.
- 55
- Abb. 7**
- [0082] Vorbereiten von (pRTA.73, BamHIA) auf die direkte Klonierung der synthetischen Spidroingene aus p9905 oder p9609 – Aufheben der SmaI-Erkennungssequenz (Position 463).
- 60
- Abb. 8**
- [0083] Einführung der Restriktionserkennungssequenzen von SmaI und NaeI in den Vektor (pRTA.73, BamHIA+SmaI) für die Klonierung synthetischer Spidroingene.
- 65
- Abb. 9**
- [0084] Allgemeine Darstellung der Herstellung transgener Spinnseidenprotein-produzierender Pflanzen.

## Abb. 10

- [0085] (a) Darstellung des modulhaften Aufbaus der erfindungsgemäßen Spinnenseidenproteine am Beispiel der SO1-Sequenz. Aminosäuren 1–28: LeB4-Signalpeptid; Aminosäuren 29–659: synthetische Spinnenseidenproteinsequenz; Aminosäuren 660–672: c-myc-Tag; Aminosäuren 673–676 ER-Retentionssignal.
- [0086] Anordnung der Sequenzmodule nach der in Simmons et al., 1996. Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. Science 271: 84–87 angegebenen Originalsequenz.
- [0087] (b) Darstellung des modulhaften Aufbaus des synthetischen Faserhybridproteins FA2. Aminosäuren 1–28: LeB4-Signalpeptid; Aminosäuren 28–130: synthetische Faserproteinsequenz der Spinne; Aminosäuren 131–247: synthetische Faserproteinsequenz des Seidenspinners; Aminosäuren 248–260: c-myc-Tag; Aminosäuren 261–264: ER-Retentionssignal.

## Abb. 11

- [0088] Schematische Darstellung der Erstellung von Genkassetten für die Akkumulation von synthetischen Faserproteinen der Spinne und des Seidenspinners im ER von transgenen Pflanzen.

## Abb. 12

- [0089] (a) Expression der synthetischen Faserproteine der Spinne (SD1, SM12, SO1) bzw. des Hybrids aus Spinne und Seidenspinner (FA2) in Blättern von transgenen Tabakpflanzen. Analysiert wurden jeweils 40 µg Gesamtprotein in SDS-Probenpuffer. SD1: 13 kDa; FA2: 20 kDa; SM12: 31 kDa; SO1: 52 kDa; K: Positivkontrolle 50 ng ScFv.
- [0090] (b) Expression der synthetischen Faserproteine der Spinne (SD1, SM12, SO1) bzw. des Hybrids aus Spinne und Seidenspinner (FA2) in transgenen Kartoffelpflanzen.
- [0091] Analysiert wurden ebenfalls jeweils 40 µg Gesamtprotein in SDS-Probenpuffer. SD1: 13 kDa; FA2: 20 kDa; SM12: 31 kDa; SO1: 52 kDa; K: Positivkontrolle 50 ng ScFv.

## Abb. 13

- [0092] Darstellung der Hitzebeständigkeit der synthetischen Faserproteine der Spinne und des Seidenspinners anhand der Konstrukte SD1 und FA2. A: Coomassie-gefärbtes Gel. B: Immunochemischer Nachweis der synthetischen Faserproteine SD1 und FA2 mittels Anti-c-myc-Antikörper. PM: Proteinmarker; ScFv: 50 ng ScFv; R: wäbriges Pflanzenextrakt von Blättern von transgenen Pflanzen für SD1 und FA2; H: Hitzeschritt 60 min, 120 min, 180 min, 24 h und 48 h bei 90°C.
- [0093] Bei der Hitzebehandlung ausgefallene Pflanzenextraktbestandteile wurden durch Zentrifugieren abgetrennt.

## Abb. 14

- [0094] Untersuchung der Lösungseigenschaften und Stabilität des synthetischen Spinnenseidenproteins SO1 nach Ammoniumsulfatfällung.
- [0095] 10 g Blattmaterial wurden in Stickstoff schockgefroren, zermörsert, in 20 ml Rohextraktpuffer aufgenommen, für 30 min bei 38°C geschüttelt, und unlösliche Bestandteile wurden durch Zentrifugieren entfernt (30 min. 10.000 rpm). Anschließend wurde der Überstand (R) für 10 min auf 90°C erhitzt und das Präzipitat durch Zentrifugieren entfernt (30 min. 10.000 rpm). Der Überstand (H) wurde mit 20% Ammoniumsulfat versetzt, 4 h bei Raumtemperatur gerollt und das Präzipitat durch Zentrifugieren für 60 min bei 4000 rpm und 4°C entfernt. Der Überstand wurde auf 30% Endkonzentration Ammoniumsulfat eingestellt und über Nacht bei Raumtemperatur gerollt. Die Lösung wurde in 5 Aliquote getrennt und das Präzipitat durch Zentrifugieren entfernt (60 min. 4000 rpm, 4°C). Die Überstände wurden verworfen und die verbliebenen Pellets in folgenden Lösungen aufgenommen: R1: Rohextraktpuffer (50 mM Tris/HCl pH 8,0; 100 mM NaCl, 10 mM MgSO<sub>4</sub>); S: SDS-Probenpuffer; G: 0,1 M Phosphatpuffer, 0,01 M Tris/HCl, 6 M Guanidiniumhydrochlorid/HCl pH 6,5; T: 1 × PBS, 1% TritonX-100; L: LiBr.
- [0096] Die Ansätze wurden für 1 h bei 37°C geschüttelt, und durch Zentrifugieren wurden unlösliche Bestandteile entfernt (30 min. 10.000 rpm). Anschließend wurde ein Aliquot jedes Ansatzes entnommen und für die SDS-Gelelektrophorese vorbereitet (R1, S1, G1, T1, L1). Die Ansätze wurden nun 36 h bei Raumtemperatur stehen gelassen. Durch Zentrifugieren wurden unlösliche Bestandteile entfernt (30 min, 10.000 rpm). Wiederum wurde ein Aliquot jedes Ansatzes entnommen und für die SDS-Gelelektrophorese vorbereitet (R2, S2, G2, T2, L2). Es wurden jeweils vergleichbare Volumina analysiert.

## Abb. 15

- [0097] Schematische Darstellung der Konstruktion von Genkassetten für die Akkumulation von Plasmalemma-ständigen synthetischen Faserproteinen der Spinne und des Seidenspinners in transgenen Pflanzen.

## Abb. 16

- [0098] Expression der Faserfusionsproteine SM12-HOOK, SO1-HOOK und FA2-HOOK in Blättern von transgenen Kartoffelpflanzen.

## Abb. 17

[0099] Gelelektrophoretische Analyse der Anreicherung von bakteriell exprimierten Spinnenseidenproteinen nach Fusion mit ELP's. Spinnenseidenprotein: 30.000 Dalton.

5

## Abb. 18

[0100] Western Blot-Analyse der Expression von Spinnenseiden-ELP-Fusionsproteinen in transgenen Tabakpflanzen. Jeweils 2,5 µg Gesamtpflanzenprotein wurden getrennt und die Spinnenseidenproteine auf dem Western Blot durch ECL nachgewiesen. Durch Vergleich mit dem Standard wird die Spinnenseidenproteinkonzentration auf mindestens 4% des gesamtlöslichen Proteins geschätzt.

10

## Abb. 19

[0101] DNA-Sequenz von SM12-70xELP als Pflanzenexpressionskassette.

15

## Abb. 20

[0102] Proteinsequenz von SM12-70xELP aus pflanzlicher Expression (SM12, c-myc-Tag, 70xELP, KDEL – jeweils durch Absatz gekennzeichnet).

20

## Abb. 21

[0103] DNA-Sequenz von SM12-70xELP als Expressionskassette für E. coli.

25

## Abb. 22

[0104] Proteinsequenz von SM12-70xELP aus bakterieller Expression (SM12, c-myc-Tag, 70xELP, c-myc-Tag, HisTag – jeweils durch Absatz gekennzeichnet).

30

35

40

45

50

55

60

65

# DE 101 13 781 A 1

## SEQUENZPROTOKOLL

<110> IPK - Institut für Pflanzengenetik und Kulturpflanz

5 <120> Synthetische Spinnenseidenproteine und deren Expression  
in transgenen Pflanzen

10 <130> I 7222

<140>  
<141>

15 <150> DE 100 28 212.1  
<151> 2000-06-09

20 <150> DE 100 53 478.3  
<151> 2000-10-24

25 <160> 51

<170> PatentIn Ver. 2.1

30 <210> 1  
<211> 22  
<212> DNA  
35 <213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen Sequenz:repetitive  
40 Einheit aus Spidroin-Proteinen

<400> 1  
45 tatgagcgct cccgggcagg gt 22

<210> 2  
50 <211> 38  
<212> DNA  
<213> Künstliche Sequenz

55 <220>  
<223> Beschreibung der künstlichen Sequenz:repetitive  
Einheit aus Spidroin-Proteinen

60 <400> 2  
agcttttagg taccaatatt aatctggccg gctccacc 38

65

# DE 101 13 781 A 1

<210> 3		
<211> 12		
<212> DNA		5
<213> Künstliche Sequenz		
<220>		
<223> Beschreibung der künstlichen Sequenz:repetitive Einheit aus Spidroin-Proteinen		10
<400> 3		
tatggtcttg gg	12	15
<210> 4		20
<211> 18		
<212> DNA		
<213> Künstliche Sequenz		25
<220>		
<223> Beschreibung der künstlichen Sequenz:repetitive Einheit aus Spidroin-Proteinen		30
<400> 4		
ggccagggtg ctggccaa	18	35
<210> 5		
<211> 33		40
<212> DNA		
<213> Künstliche Sequenz		
<220>		45
<223> Beschreibung der künstlichen Sequenz:repetitive Einheit aus Spidroin-Proteinen		
<400> 5		50
ggtgcaggag cwgcwgcwgc wgctgcaggt gga	33	
<210> 6		55
<211> 28		
<212> DNA		
<213> Künstliche Sequenz		60
<220>		
<223> Beschreibung der künstlichen Sequenz:repetitive Einheit aus Spidroin-Proteinen		65

<400> 6  
 gccggccaga ttaatattgg tacctaaa 28  
 5  
 <210> 7  
 <211> 17  
 10 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 15 <223> Beschreibung der künstlichen Sequenz:repetitive  
 Einheit aus Spidroin-Proteinen  
 20 <400> 7  
 ctgcccggga gcgctca 17  
 25 <210> 8  
 <211> 15  
 <212> DNA  
 30 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:repetitive  
 35 Einheit aus Spidroin-Proteinen  
 <400> 8  
 accaccataa cctcc 15  
 40  
 <210> 9  
 45 <211> 18  
 <212> DNA  
 <213> Künstliche Sequenz  
 50 <220>  
 <223> Beschreibung der künstlichen Sequenz:repetitive  
 Einheit aus Spidroin-Proteinen  
 55 <400> 9  
 agcaccctgg cccccag 18  
 60  
 <210> 10  
 <211> 33  
 <212> DNA  
 65 <213> Künstliche Sequenz

# DE 101 13 781 A 1

<220>		
<223> Beschreibung der künstlichen Sequenz:repetitive Einheit aus Spidroin-Proteinen		5
<400> 10 tgcagcwgcw gcwgcgctc ctgcaccttg gcc	33	10
<210> 11 <211> 22 <212> DNA <213> Künstliche Sequenz		15
<220> <223> Beschreibung der künstlichen Sequenz:repetitive Einheit aus Spidroin-Proteinen		20
<400> 11 tatgagatct ggccaaggag gt	22	25
<210> 12 <211> 14 <212> DNA <213> Künstliche Sequenz		30
<220> <223> Beschreibung der künstlichen Sequenz:repetitive Einheit aus Spidroin-Proteinen		35
<400> 12 ttggccagat ctca	14	40
<210> 13 <211> 27 <212> DNA <213> Künstliche Sequenz		45
<220> <223> Beschreibung der künstlichen Sequenz:repetitive Einheit aus Spidroin-Proteinen		50
<400> 13 agtcagggtg ctggtcgtgg aggccaa	27	55
<210> 14		60
		65



# DE 101 13 781 A 1

<211> 27  
 <212> DNA  
 <213> Künstliche Sequenz  
 5  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:repetitive  
 10 Einheit aus Spidroin-Proteinen  
 <400> 14  
 15 tccacgacca gcaccctgac tccccag 27  
 <210> 15  
 20 <211> 36  
 <212> DNA  
 <213> Künstliche Sequenz  
 25 <220>  
 <223> Beschreibung der künstlichen Sequenz:repetitive  
 Einheit aus Spidroin-Proteinen  
 30 <400> 15  
 agtcagggcg ctggtcgtgg gggactgggt ggccaa 36  
 35 <210> 16  
 <211> 36  
 <212> DNA  
 40 <213> Künstliche Sequenz  
 <220>  
 45 <223> Beschreibung der künstlichen Sequenz:repetitive  
 Einheit aus Spidroin-Proteinen  
 <400> 16  
 50 acccagtcac ccacgaccag cgcctgact cccag 36  
 <210> 17  
 55 <211> 24  
 <212> DNA  
 <213> Künstliche Sequenz  
 60 <220>  
 <223> Beschreibung der künstlichen Sequenz:repetitive  
 Einheit aus Spidroin-Proteinen  
 65 <400> 17

# DE 101 13 781 A 1

ctgggagggc agggagcggg ccaa

24

<210> 18

5

<211> 24

<212> DNA

<213> Künstliche Sequenz

10

<220>

<223> Beschreibung der künstlichen Sequenz:repetitive  
Einheit aus Spidroin-Proteinen

15

<400> 18

cgctccctgc cctcccagac ctcc

24

20

<210> 19

<211> 327

25

<212> DNA

<213> Künstliche Sequenz

30

<220>

<223> Beschreibung der künstlichen Sequenz:Konstrukt SB1

<400> 19

35

ggatcccagt tagggcaggg aggttatggt ggtctggggg gccaggggtgc tggccaagga 60  
ggttatggtg gtctggggag tcagggcgct ggtcgtgggg gactgggtgg ccaaggtgca 120  
ggagctgctg ctgcagctgc aggtggagcc gggcagggag gtctgggagg gcagggagcg 180  
ggccaaggtg caggagcagc tgcagcagct gcaggtggag ccgggcaggg aggttatggt 240  
ggtctgggga gtcaggggtgc tggtcgtgga ggccaaggtg caggagctgc agcagcagct 300  
gcaggtggag ccggacaagc ggccgca 327

40

45

<210> 20

<211> 705

<212> DNA

50

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Konstrukt SE1

55

<400> 20

ggatcccagt tagggcaggg aggttatggt ggtctggggg gccaggggtgc tggccaagga 60  
ggttatggtg gtctggggag tcagggcgct ggtcgtgggg gactgggtgg ccaaggtgca 120  
ggagctgctg ctgcagctgc aggtggagcc gggcagggag gtctgggagg gcagggagcg 180  
ggccaaggtg caggagcagc tgcagcagct gcaggtggag ccgggcaggg aggttatggt 240  
ggtctgggga gtcagggcgct tggtcgtggg ggactgggtg gccaaaggtgc aggagcagct 300  
gcagctgctg caggtggagc cgggcagggg ggttatggtg gtctggggag tcaggggtgct 360

65

# DE 101 13 781 A 1

ggctcgtggag gccaaaggtgc aggagctgca gcagcagctg caggtggagc cgggcagggg 420  
 gggttatggtg gtctggggag tcagggcgct ggtcgtgggg gactgggtgg ccaaggtgca 480  
 ggagcagctg cagctgctgc aggtggagcc gggcagggag gttatggtgg tctggggagt 540  
 5 cagggtgctg gtcgtggagg ccaaggtgca ggagctgcag cagcagctgc aggtggagcc 600  
 gggcagggag gttatggtgg tctggggagt cagggtgctg gtcgtggagg ccaaggtgca 660  
 ggagctgcag cagcagctgc aggtggagcc ggacaagcgg ccgca 705

10

<210> 21

<211> 426

15

<212> DNA

<213> Künstliche Sequenz

20 <220>

<223> Beschreibung der künstlichen Sequenz:Konstrukt SD1

<400> 21

25 ggatcccagt tagggcaggg aggttatggt ggtctggggg gccaggggtgc tggccaagga 60  
 gggttatggtg gtctggggag tcagggcgct ggtcgtgggg gactgggtgg ccaaggtgca 120  
 ggagctgctg ctgcagctgc aggtggagcc gggcagggag gtctgggagg gcagggagcg 180  
 ggccaaggtg caggagcagc tgcagcagct gcaggtggag ccgggcaggg aggttatggt 240  
 30 ggtctgggga gtcaggggtg tggctgtgga ggccaaggtg caggagctgc agcagcagct 300  
 gcaggtggag ccgggcaggg aggttatggt ggtctgggga gtcagggcgc tggctgtggg 360  
 ggactgggtg gccaaaggtgc aggagcagct gcagctgctg caggtggagc cggacaagcg 420  
 35 gccgca 426

<210> 22

40

<211> 3783

<212> DNA

<213> Künstliche Sequenz

45

<220>

<223> Beschreibung der künstlichen Sequenz:Konstrukt  
 S01S01

50

<400> 22

ggatcccagt taccggggca gggaggttat ggtggtctgg ggggccaggg tgctggccaa 60  
 ggaggttatg gtggtctggg gggccaggg gctggccaag gtgcaggagc tgctgctgca 120  
 55 gctgcagggtg gagccgggca gggaggttat ggtggtctgg ggagtcaggg tgctggtcgt 180  
 ggaggccaag gtgcaggagc tgcagcagca gctgcagggtg gagccgggca gggaggttat 240  
 ggtggtctgg ggagtcaggg cgtggtcgt gggggactgg gtggccaagg tgcaggagca 300  
 60 gctgcagctg ctgcaggtgg agccgggcag ggaggttatg gtggtctggg gagtcagggt 360  
 gctggtcgtg gagccaagg tgcaggagct gcagcagcag ctgcaggtgg agccgggcag 420  
 ggaggttatg gtggtctggg ggtcagggc gctggtcgtg ggggactggg tggccaaggt 480  
 gcaggagcag ctgcagctgc tgcaggtgga gccgggcagg gaggttatgg tggctctgggg 540  
 65 ggccaggggtg ctggccaagg aggttatggt ggtctgggga gtcagggcgc tggctgtggg 600  
 ggactgggtg gccaaaggtgc aggagctgct gctgcagctg caggtggagc cgggcagggg 660

# DE 101 13 781 A 1

ggctctgggag	ggcagggagc	gggccaaggt	gcaggagcag	ctgcagcagc	tgcaggtgga	720	
gccgggcagg	gaggttatgg	tggctctggg	agtcaggggtg	ctggctcgtgg	aggccaaggt	780	
gcaggagctg	cagcagcagc	tgcaggtgga	gccgggcagg	gaggttatgg	tggctctggg	840	5
ggccaaggtg	ctggccaagg	aggttatggt	ggctctgggga	gtcagggcgc	tggctcgtggg	900	
ggactgggtg	gccaaggtgc	aggagctgct	gctgcagctg	caggtggagc	cgggcaggga	960	
ggctctgggag	ggcagggagc	gggccaaggt	gcaggagcag	ctgcagcagc	tgcaggtgga	1020	
gccgggcagg	gaggttatgg	tggctctggg	agtcagggcg	ctggctcgtgg	gggactgggt	1080	10
ggccaaggtg	caggagcagc	tgcagctgct	gcaggtggag	cggggcaggg	aggttatggt	1140	
ggctctgggga	gtcaggggtgc	tggctcgtgga	ggccaaggtg	caggagctgc	agcagcagct	1200	
gcaggtggag	cggggcaggg	aggttatggt	ggctctgggga	gtcagggcgc	tggctcgtggg	1260	
ggactgggtg	gccaaggtgc	aggagcagct	gcagctgctg	caggtggagc	cgggcaggga	1320	15
ggttatgggtg	gtctggggag	tcaggttgct	ggctcgtggag	gccaaggtgc	aggagctgca	1380	
gcagcagctg	caggtggagc	cgggcaggga	ggttatgggtg	gtctggggag	tcaggggtgct	1440	
ggctcgtggag	gccaaggtgc	aggagctgca	gcagcagctg	caggtggagc	cgggcaggga	1500	20
ggttatgggtg	gtctgggggg	ccaggttgct	ggccaaggag	gttatgggtg	tctggggagt	1560	
cagggcgctg	gtcgtggggg	actgggtggc	caaggtgcag	gagctgctgc	tgcagctgca	1620	
ggtggagccg	ggcagggagg	tctgggagg	cagggagcgg	gccaaggtgc	aggagcagct	1680	
gcagcagctg	caggtggagc	cgggcaggga	ggttatgggtg	gtctggggag	tcaggggtgct	1740	25
ggctcgtggag	gccaaggtgc	aggagctgca	gcagcagctg	caggtggagc	cgggcaggga	1800	
ggttatgggtg	gtctggggag	tcagggcgct	ggctcgtggg	gactgggtgg	ccaaggtgca	1860	
ggagcagctg	cagctgctgc	agggtggagcc	gggcagggag	gttatgggtg	tctggggggc	1920	30
caggggtgctg	gccaaggagg	ttatgggtgg	ctggggggcc	agggtgctgg	ccaaggtgca	1980	
ggagctgctg	ctgcagctgc	agggtggagcc	gggcagggag	gttatgggtg	tctggggagt	2040	
caggggtgctg	gtcgtggagg	ccaaggtgca	ggagctgcag	cagcagctgc	agggtggagcc	2100	
gggcagggag	gttatgggtg	tctggggagt	cagggcgctg	gtcgtggggg	actgggtggc	2160	35
caaggtgcag	gagcagctgc	agctgctgca	ggtggagccg	ggcagggagg	ttatgggtgg	2220	
ctggggagtc	agggtgctgg	tcgtggaggc	caaggtgcag	gagctgcagc	agcagctgca	2280	
ggtggagccg	ggcagggagg	ttatgggtgg	ctggggagtc	agggcgctgg	tcgtggggga	2340	
ctgggtggcc	aaggtgcagg	agcagctgca	gctgctgcag	gtggagccgg	gcagggaggt	2400	40
tatgggtggtc	tggggggcca	gggtgctggc	caaggaggtt	atgggtggtc	ggggagtcag	2460	
ggcgctgggtc	gtgggggact	gggtggccaa	ggtgcaggag	ctgctgctgc	agctgcaggt	2520	
ggagccggggc	agggaggtct	gggagggcag	ggagccgggc	aaggtgcagg	agcagctgca	2580	45
gcagctgcag	gtggagccgg	gcagggaggt	tatgggtggtc	tggggagtca	gggtgctggt	2640	
cgtggaggcc	aaggtgcagg	agctgcagca	gcagctgcag	gtggagccgg	gcagggaggt	2700	
tatgggtggtc	tggggggcca	gggtgctggc	caaggaggtt	atgggtggtc	ggggagtcag	2760	
ggcgctgggtc	gtgggggact	gggtggccaa	ggtgcaggag	ctgctgctgc	agctgcaggt	2820	50
ggagccggggc	agggaggtct	gggagggcag	ggagccgggc	aaggtgcagg	agcagctgca	2880	
gcagctgcag	gtggagccgg	gcagggaggt	tatgggtggtc	tggggagtca	gggcgctggt	2940	
cgtgggggac	tgggtggcca	agggtgcagg	gcagctgcag	ctgctgcagg	tggagccggg	3000	55
cagggaggtt	atgggtggtc	ggggagtcag	ggtgctggtc	gtggaggcca	agggtgcagga	3060	
gctgcagcag	cagctgcagg	tggagccggg	cagggaggtt	atgggtggtc	ggggagtcag	3120	
ggcgctgggtc	gtgggggact	gggtggccaa	ggtgcaggag	cagctgcagc	tgctgcaggt	3180	
ggagccggggc	agggaggtta	tgggtggtctg	gggagtcagg	gtgctggtcg	tggaggccaa	3240	60
ggtgcaggag	ctgcagcagc	agctgcaggt	ggagccgggc	agggaggtta	tgggtggtctg	3300	
gggagtcagg	gtgctggtcg	tggaggccaa	ggtgcaggag	ctgcagcagc	agctgcaggt	3360	
ggagccggggc	agggaggtta	tgggtggtctg	ggggggcagg	gtgctggcca	aggaggttat	3420	65
ggtggtctgg	ggagtcaggg	cgctggtcgt	gggggactgg	gtggccaagg	tgcaggagct	3480	
gctgctgcag	ctgcaggtgg	agccgggcag	ggaggtctgg	gagggcaggg	agcggggccaa	3540	

# DE 101 13 781 A 1

ggtgcaggag cagctgcagc agctgcaggt ggagccgggc agggaggtta tgggtggtctg 3600  
 gggagtcagg gtgctggtcg tggaggccaa ggtgcaggag ctgcagcagc agctgcaggt 3660  
 gggagccgggc agggaggtta tgggtggtctg gggagtcagg gcgctggtcg tgggggactg 3720  
 5 ggtggccaaag gtgcaggagc agctgcagct gctgcaggtg gagccggcgg acaagcggcc 3780  
 gca 3783

10

<210> 23

<211> 2985

<212> DNA

15

<213> Künstliche Sequenz

<220>

20

<223> Beschreibung der künstlichen Sequenz:Konstrukt  
 S01SM12

<400> 23

25

ggatcccagt taccggggca gggaggttat ggtggtctg ggggccaggg tgctggccaa 60  
 gggaggttat gtggtctggg gggccagggt gctggccaag gtgcaggagc tgctgctgca 120  
 gctgcaggtg gagccgggca gggaggttat ggtggtctg ggagtcaggg tgctggtcgt 180  
 gggagccaaag gtgcaggagc tgcagcagca gctgcaggtg gagccgggca gggaggttat 240  
 30 ggtggtctg gggagtcagg cgctggtcgt gggggactgg gtggccaagg tgcaggagca 300  
 gctgcagctg ctgcaggtgg agccgggcag ggaggttat gtggtctggg gagtcagggt 360  
 gctggtcgtg gagccaagg tgcaggagct gcagcagcag ctgcaggtgg agccgggcag 420  
 35 gggaggttat gtggtctggg gagtcagggt gctggtcgt ggggactggg tggccaagg 480  
 gcaggagcag ctgcagctgc tgcaggtgga gccgggcagg gaggttatgg tggctctggg 540  
 ggccagggtg ctggccaagg aggttatggt ggtctgggga gtcagggcgc tggctcgtgg 600  
 ggactgggtg gccaagggtg aggagctgct gctgcagctg cagggtggag cgggcaggga 660  
 40 ggtctgggag ggcagggagc gggccaagg gtgcaggagc ctgcagcagc tgcagggtgga 720  
 gccgggcagg gaggttatgg tggctctggg agtcagggtg ctggtcgtgg aggccaagg 780  
 gcaggagctg cagcagcagc tgcaggtgga gccgggcagg gaggttatgg tggctctggg 840  
 45 ggccagggtg ctggccaagg aggttatggt ggtctgggga gtcagggcgc tggctcgtgg 900  
 ggactgggtg gccaagggtg aggagctgct gctgcagctg cagggtggag cgggcaggga 960  
 ggtctgggag ggcagggagc gggccaagg gtgcaggagc ctgcagcagc tgcagggtgga 1020  
 gccgggcagg gaggttatgg tggctctggg agtcagggcg ctggtcgtgg gggactgggt 1080  
 50 ggccaagggt caggagcagc tgcagctgct gcagggtggag ccgggcaggg aggttatggt 1140  
 ggtctgggga gtcagggtgc tggctcgtgga ggccaagggt caggagctgc agcagcagct 1200  
 gcagggtggag ccgggcaggg aggttatggt ggtctgggga gtcagggcgc tggctcgtgg 1260  
 ggactgggtg gccaagggtg aggagcagct gcagctgctg cagggtggag cgggcaggga 1320  
 55 ggttatggtg gtctggggag tcagggtgct ggtcgtggag gccaagggtg aggagctgca 1380  
 gcagcagctg cagggtggag cgggcaggga ggttatggt gtctggggag tcagggtgct 1440  
 ggtcgtggag gccaagggtg aggagctgca gcagcagctg cagggtggag cgggcaggga 1500  
 60 ggttatggtg gtctgggggg ccagggtgct ggccaaggag gttatggtgg tctggggagt 1560  
 cagggcgctg gtcgtggggg a>tgggtggc caagggtgag gagctgctgc tgcagctgca 1620  
 ggtggagccg ggcagggagg tctgggaggg caggagcgg gccaagggtg aggagcagct 1680  
 gcagcagctg cagggtggag cgggcaggga ggttatggt gtctggggag tcagggtgct 1740  
 65 ggtcgtggag gccaagggtg aggagctgca gcagcagctg cagggtggag cgggcaggga 1800  
 ggttatggtg gtctggggag tcagggcgct ggtcgtgggg gactgggtgg ccaagggtgca 1860

# DE 101 13 781 A 1

ggagcagctg	cagctgctgc	aggtggagcc	gggcagggag	gttatggtgg	tctggggggc	1920	
cagggtgctg	gccaaaggagg	ttatggtggt	ctggggagtc	agggcgctgg	tcgtggggga	1980	
ctgggtggcc	aaggtgcagg	agctgctgct	gcagctgcag	gtggagccgg	gcagggaggt	2040	
ctgggagggc	agggagcggg	ccaaggtgca	ggagcagctg	cagcagctgc	aggtggagcc	2100	5
gggcagggag	gttatggtgg	tctggggagt	cagggcgctg	gtcgtggggg	actgggtggc	2160	
caaggtgcag	gagcagctgc	agctgctgca	ggtggagccg	ggcagggagg	ttatggtggt	2220	
ctggggagtc	aggtgctggg	tcgtggaggg	caaggtgcag	gagctgcagc	agcagctgca	2280	10
ggtggagccg	ggcagggagg	ttatggtggt	ctggggagtc	agggcgctgg	tcgtggggga	2340	
ctgggtggcc	aaggtgcagg	agcagctgca	gctgctgcag	gtggagccgg	gcagggaggt	2400	
tatggtggtc	tggggagtca	gggtgctggt	cgtggaggcc	aaggtgcagg	agctgcagca	2460	
gcagctgcag	gtggagccgg	gcagggaggt	tatggtggtc	tggggagtca	gggtgctggt	2520	15
cgtggaggcc	aaggtgcagg	agctgcagca	gcagctgcag	gtggagccgg	gcagggaggt	2580	
tatggtggtc	tggggggcca	gggtgctggc	caaggaggtt	atggtggtct	ggggagttag	2640	
ggcgctggtc	gtgggggact	gggtggccaa	ggtgcaggag	ctgctgctgc	agctgcaggt	2700	20
ggagccgggc	agggaggtct	gggagggcag	ggagcggggc	aaggtgcagg	agcagctgca	2760	
gcagctgcag	gtggagccgg	gcagggaggt	tatggtggtc	tggggagtca	gggtgctggt	2820	
cgtggaggcc	aaggtgcagg	agctgcagca	gcagctgcag	gtggagccgg	gcagggaggt	2880	
tatggtggtc	tggggagtca	gggcgctggt	cgtgggggac	tgggtggcca	aggtgcagga	2940	25
gcagctgcag	ctgctgcagg	tggagccggc	ggacaagcgg	ccgca		2985	
							30
<210> 24							
<211> 5658							
<212> DNA							
<213> Künstliche Sequenz							35
<220>							
<223> Beschreibung der künstlichen Sequenz:Konstrukt							
S01S01S01							40
<400> 24							
ggatcccagt	taccggggca	gggaggttat	ggtggctctg	ggggccaggg	tgctggccaa	60	45
ggaggttatg	gtggtctggg	gggcccagggt	gctggccaag	gtgcaggagc	tgctgctgca	120	
gctgcagggtg	gagccggggca	gggaggttat	ggtggctctg	ggagtcaggg	tgctggtcgt	180	
ggaggccaag	gtgcaggagc	tgagcagca	gctgcagggtg	gagccggggca	gggaggttat	240	
ggtggtctg	ggagtcaggg	cgctggtcgt	gggggactgg	gtggccaagg	tgagggagca	300	50
gctgcagctg	ctgcagggtg	agccggggcag	ggaggttatg	gtggtctggg	gagtcagggt	360	
gctggtcgtg	gaggccaagg	tgagggagct	gcagcagcag	ctgcagggtg	agccggggcag	420	
ggaggttatg	gtggtctggg	gagtcagggt	gctggtcgtg	ggggactggg	tggccaagggt	480	
gcagggagcag	ctgcagctgc	tgaggtgga	gccggggcagg	gaggttatgg	tgggtctgggg	540	55
ggccagggtg	ctggccaagg	aggttatggt	ggtctggggg	gtcagggcgc	tgggtcgtggg	600	
ggactgggtg	gccaagggtg	aggagctgct	gctgcagctg	caggtggagc	cgggcaggga	660	
ggtctggggg	ggcagggagc	gggccaagggt	gcaggagcag	ctgcagcagc	tgaggtgga	720	60
gccggggcagg	gaggttatgg	tgggtctgggg	agtcagggtg	ctggtcgtgg	aggccaagggt	780	
gcagggagctg	cagcagcagc	tgaggtgga	gccggggcagg	gaggttatgg	tgggtctgggg	840	
ggccagggtg	ctggccaagg	aggttatggt	ggtctggggg	gtcagggcgc	tgggtcgtggg	900	
ggactgggtg	gccaagggtg	aggagctgct	gctgcagctg	caggtggagc	cgggcaggga	960	65
ggtctggggg	ggcagggagc	gggccaagggt	gcaggagcag	ctgcagcagc	tgaggtgga	1020	

# DE 101 13 781 A 1

gccgggcagg gaggttatgg tggctctgggg agtcagggcg ctggctcgtgg gggactgggt 1080  
 ggccaagggtg caggagcagc tgcagctgct gcagggtggag ccggggcaggg aggttatgggt 1140  
 5 ggtctggggga gtcagggtgc tggctcgtgga ggccaagggtg caggagctgc agcagcagct 1200  
 gcagggtggag ccggggcaggg aggttatgggt ggtctggggga gtcaggggcgc tggctcgtggg 1260  
 ggactgggtg gccaaagggtgc aggagcagct gcagctgctg cagggtggagc cggggcaggga 1320  
 gggttatgggtg gtctggggag tcagggtgct ggtcgtggag gccaaagggtgc aggagctgca 1380  
 10 gcagcagctg cagggtggagc cggggcaggga gggttatgggtg gtctggggag tcagggtgct 1440  
 ggtcgtggag gccaaagggtgc aggagctgca gcagcagctg cagggtggagc cggggcaggga 1500  
 gggttatgggtg gtctgggggg ccagggtgct ggccaaggag gttatgggtgg tctggggagt 1560  
 caggggcgtg gtcgtggggg actgggtggc caagggtgcag gagctgctgc tgcagctgca 1620  
 15 ggtggagccg ggcaaggaggg tctgggaggg caggggagcgg gccaaagggtgc aggagcagct 1680  
 gcagcagctg cagggtggagc cggggcaggga gggttatgggtg gtctggggag tcagggtgct 1740  
 ggtcgtggag gccaaagggtgc aggagctgca gcagcagctg cagggtggagc cggggcaggga 1800  
 20 gggttatgggtg gtctggggag tcaggggcgt ggtcgtgggg gactgggtgg ccaagggtgca 1860  
 ggagcagctg cagctgctgc aggtggagcc gggcaggggag gttatgggtgg tctggggggc 1920  
 cagggtgctg gccaaaggag ttatgggtgg ctggggggcc aggggtgctgg ccaagggtgca 1980  
 ggagctgctg ctgcagctgc aggtggagcc gggcaggggag gttatgggtgg tctggggagt 2040  
 25 cagggtgctg gtcgtggagg ccaagggtgca ggagctgcag cagcagctgc aggtggagcc 2100  
 gggcaggggag gttatgggtgg tctggggagt caggggcgtg gtcgtggggg actgggtggc 2160  
 caagggtgcag gagcagctgc agctgctgca ggtggagccg ggcaaggaggg ttatgggtgg 2220  
 ctggggagtc aggggtgctgg tcgtggaggc caagggtgcag gagctgcagc agcagctgca 2280  
 30 ggtggagccg ggcaaggaggg ttatgggtgg ctggggagtc agggcgctgg tcgtggggga 2340  
 ctgggtggcc aagggtgcagg agcagctgca gctgctgcag gtggagccgg gcaggggaggt 2400  
 tatgggtggtc tggggggcca ggggtgctggc caaggaggtt atgggtggtct ggggagtcag 2460  
 35 ggcgctggtc gtgggggact ggggtggcca ggtgcaggag ctgctgctgc agctgcaggt 2520  
 ggagccgggc agggaggtct gggagggcag ggagccggcc aagggtgcagg agcagctgca 2580  
 gcagctgcag gtggagccgg gcaggggaggt tatgggtggtc tggggagtca ggggtgctgg 2640  
 cgtggaggcc aagggtgcagg agctgcagca gcagctgcag gtggagccgg gcaggggaggt 2700  
 40 tatgggtggtc tggggggcca ggggtgctggc caaggaggtt atgggtggtct ggggagtcag 2760  
 ggcgctggtc gtgggggact ggggtggcca ggtgcaggag ctgctgctgc agctgcaggt 2820  
 ggagccgggc agggaggtct gggagggcag ggagccggcc aagggtgcagg agcagctgca 2880  
 45 gcagctgcag gtggagccgg gcaggggaggt tatgggtggtc tggggagtca gggcgctgg 2940  
 cgtgggggac tgggtggcca aggtgcaggga gcagctgcag ctgctgcagg tggagccggg 3000  
 caggggaggtt atgggtggtct ggggagtcag ggtgctggtc gtggaggcca aggtgcaggga 3060  
 gctgcagcag cagctgcagg tggagccggg caggggaggtt atgggtggtct ggggagtcag 3120  
 50 ggcgctggtc gtgggggact ggggtggcca ggtgcaggag cagctgcagc tgctgcaggt 3180  
 ggagccgggc agggaggtta tgggtggtctg gggagtcagg gtgctggtcg tggaggcca 3240  
 ggtgcaggag ctgcagcagc agctgcaggt ggagccgggc agggaggtta tgggtggtctg 3300  
 55 gggagtcagg gtgctggtcg tggaggcca ggtgcaggag ctgcagcagc agctgcaggt 3360  
 ggagccgggc agggaggtta tgggtggtctg gggggccagg gtgctggcca aggaggttat 3420  
 ggtggtctgg ggagtcaggg cgctggtcgt gggggactgg gtggccaagg tgcaggagct 3480  
 gctgctgcag ctgcaggtgg agccgggcag ggaggtctgg gagggcaggg agcggggcca 3540  
 60 ggtgcaggag cagctgcagc agctgcaggt ggagccgggc agggaggtta tgggtggtctg 3600  
 gggagtcagg gtgctggtcg tggaggcca ggtgcaggag ctgcagcagc agctgcaggt 3660  
 ggagccgggc agggaggtta tgggtggtctg gggagtcagg gcgctggtcg tgggggactg 3720  
 ggtggccaag gtgcaggagc agctgcagct gctgcaggtg gagccgggca gggaggttat 3780  
 65 ggtggtctcg ggggccaggg tgctggcca ggaggttatg gtggtctggg gggccagggt 3840  
 gctggccaag gtgcaggagc tgctgctgca gctgcaggtg gagccgggca gggaggttat 3900

# DE 101 13 781 A 1

ggtggtcttg ggagtcaggg tgctggtcgt ggaggccaag gtgcaggagc tgcagcagca 3960  
 gctgcagggtg gagccgggca gggaggttat ggtggtcttg ggagtcaggg cgctggtcgt 4020  
 gggggacttg gtggccaagg tgcaggagca gctgcagctg ctgcagggtg agccgggcag 4080  
 ggaggttatg gtggtctggg gagtcagggt gctggtcgtg gaggccaagg tgcaggagct 4140  
 gcagcagcag ctgcagggtg agccgggcag ggaggttatg gtggtctggg gagtcagggtc 4200  
 gctggtcgtg ggggactggg tggccaaggc gcaggagcag ctgcagctgc tgcagggtgga 4260  
 gccgggcagg gaggttatgg tggctctggg ggcagggtg ctggccaagg aggttatggg 4320  
 ggtctgggga gtcaggggcg tggctctggg ggactgggtg gccaagggtg aggagctgct 4380  
 gctgcagctg cagggtggag cgggcaggga ggtctgggag ggcagggtg gggccaaggc 4440  
 gcaggagcag ctgcagcagc tgcagggtgga gccgggcagg gaggttatgg tggctctggg 4500  
 agtcagggtg ctggtcgtgg aggccaaggc gcaggagctg cagcagcagc tgcagggtgga 4560  
 gccgggcagg gaggttatgg tggctctggg ggcagggtg ctggccaagg aggttatggg 4620  
 ggtctgggga gtcaggggcg tggctctggg ggactgggtg gccaagggtg aggagctgct 4680  
 gctgcagctg cagggtggag cgggcaggga ggtctgggag ggcagggtg gggccaaggc 4740  
 gcaggagcag ctgcagcagc tgcagggtgga gccgggcagg gaggttatgg tggctctggg 4800  
 agtcagggtg ctggtcgtgg gggactgggt ggcgaagggt caggagcagc tgcagctgct 4860  
 gcagggtggag ccgggcaggg aggttatggg ggtctgggga gtcagggtg tggctcgtgga 4920  
 ggccaagggt caggagctgc agcagcagct gcagggtggag ccgggcaggg aggttatggg 4980  
 ggtctgggga gtcaggggcg tggctctggg ggactgggtg gccaagggtg aggagcagct 5040  
 gcagctgctg cagggtggag cgggcaggga ggttatggg gtctggggag tcagggtgct 5100  
 ggtcgtggag gccaagggtg aggagctgca gcagcagctg cagggtggag cgggcaggga 5160  
 ggttatggg gtctggggag tcagggtgct ggtcgtggag gccaagggtg aggagctgca 5220  
 gcagcagctg cagggtggag cgggcaggga ggttatggg gtctgggggg ccagggtgct 5280  
 ggccaaggag gttatggtg tctggggagt cagggtcgtg gtcgtggggg actgggtggc 5340  
 caagggtgag gagctgctg tgcagctgca ggtggagccg ggcagggtg tctgggaggg 5400  
 cagggtgagg gccaagggtg aggagcagct gcagcagctg cagggtggag cgggcaggga 5460  
 ggttatggg gtctggggag tcagggtgct ggtcgtggag gccaagggtg aggagctgca 5520  
 gcagcagctg cagggtggag cgggcaggga ggttatggg gtctggggag tcagggtgct 5580  
 ggtcgtggg gactgggtg ccaagggtgca ggagcagctg cagctgctgc aggtggagcc 5640  
 ggcggacaag cggccgca 5658

<210> 25

<211> 672

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Konstrukt FA2

<400> 25

ggatcccagt tagggcaggg aggttatggg ggtctggggg gccagggtg tggccaagga 60  
 ggttatggg gtctggggag tcagggtgct ggtcgtggg gactgggtg ccaagggtgca 120  
 ggagctgctg ctgcagctgc aggtggagcc gggcagggtg gtctgggagg gcagggtgagc 180  
 ggccaagggt caggagcagc tgcagcagct gcagggtggag ccgggcaggg aggttatggg 240  
 ggtctgggga gtcaggggcg tggctctggg ggactgggtg gccaagggtg aggagcagct 300  
 gcagctgctg cagggtggag cgggtccgga agtgggtgag gtgccggaag cggagcagga 360  
 gccggtgccg gatctggtgc cgggtgccga agcgggtgct gtgccggaag cgggtgctggg 420



# DE 101 13 781 A 1

gccggatcag gagcgggtgc cggttatggt gcgggagccg gtgttgggta cggagccggt 480  
 tatggagcgg gagccggtgt tgggtacgga gccggtgcag gttccggggc cgcaagcggc 540  
 gcaggagccg gtgccggagc tgggacaggg agttcaggat ttggggcccta cgttgcaaata 600  
 5 ggtggttatt caggctatga atacgcgtgg agtagtaagt ctgattttga gactgccgga 660  
 caagcggccg ca 672

10

<210> 26

<211> 525

<212> DNA

15

<213> Künstliche Sequenz

<220>

20

<223> Beschreibung der künstlichen Sequenz:Konstrukt SA1

<400> 26

ggatcccagt tagggcaggg aggttatggt ggtctggggg gccaggggtgc tggccaagga 60  
 25 ggttatggtg gtctgggggg ccagggtgct ggccaaggtg caggagctgc tgctgcagct 120  
 gcagggtggag ccgggcaggg aggttatggt ggtctgggga gtcaggggtgc tggctcgtgga 180  
 ggccaaggtg caggagctgc agcagcagct gcagggtggag ccgggcaggg aggttatggt 240  
 30 ggtctgggga gtcagggcgc tggctcgtgg ggactgggtg gccaaaggtg aggagcagct 300  
 gcagctgctg cagggtggagc cgggcagggg ggttatggtg gtctggggag tcagggtgct 360  
 ggtcgtggag gccaaaggtg aggagctgca gcagcagctg cagggtggagc cgggcagggg 420  
 ggttatggtg gtctggggag tcagggcgct ggtcgtgggg gactgggtgg ccaaggtgca 480  
 35 ggagcagctg cagctgctgc aggtggagcc ggacaagcgg ccgca 525

<210> 27

40

<211> 1908

<212> DNA

<213> Künstliche Sequenz

45

<220>

<223> Beschreibung der künstlichen Sequenz:Konstrukt SO1

50 <400> 27

ggatcccagt tacccgggca gggaggttat ggtggtctgg ggggccaggg tgctggccaa 60  
 ggaggttatg gtggtctggg gggccaggg gctggccaag gtgcaggagc tgctgctgca 120  
 gctgcagggtg gagccgggca gggaggttat ggtggtctgg ggagtcaggg tgctggtcgt 180  
 55 ggaggccaag gtgcaggagc tgcagcagca gctgcagggtg gagccgggca gggaggttat 240  
 ggtggtctgg ggagtcaggg cgtggtcgt gggggactgg gtggccaagg tgcaggagca 300  
 gctgcagctg ctgcagggtg agccgggcag ggaggttatg gtggtctggg gagtcagggt 360  
 60 gctggtcgtg gaggccaagg tgcaggagct gcagcagcag ctgcagggtg agccgggcag 420  
 ggaggttatg gtggtctggg gagtcagggc gctggtcgtg ggggactggg tggccaaggt 480  
 gcaggagcag ctgcagctgc tgcagggtga gccgggcagg gaggttatgg tggctctggg 540  
 ggccaggggtg ctggccaagg aggttatggt ggtctgggga gtcagggcgc tggctcgtgg 600  
 65 ggactgggtg gccaaaggtg aggagctgct gctgcagctg cagggtggagc cgggcagggg 660  
 ggtctgggag ggcaaggagc gggccaaggt gcaggagcag ctgcagcagc tgcagggtga 720

# DE 101 13 781 A 1

gccgggcagg	gaggttatgg	tggctctgggg	agtcagggtg	ctggctcgtgg	aggccaaggt	780	
gcaggagctg	cagcagcagc	tgcagggtgga	gccgggcagg	gaggttatgg	tggctctgggg	840	
ggccagggtg	ctggccaagg	aggttatggt	ggctctgggga	gtcaggggcgc	tggctcgtggg	900	5
ggactgggtg	gccaagggtgc	aggagctgct	gctgcagctg	cagggtggagc	cgggcaggga	960	
ggctctgggag	ggcaggggagc	gggccaaggt	gcaggagcag	ctgcagcagc	tgcagggtgga	1020	
gccgggcagg	gaggttatgg	tggctctgggg	agtcagggtg	ctggctcgtgg	gggactgggt	1080	
ggccaaggtg	caggagcagc	tgcagctgct	gcagggtgga	ccgggcaggg	aggttatggt	1140	10
ggctctgggga	gtcagggtgc	tggctcgtgga	ggccaaggtg	caggagctgc	agcagcagct	1200	
gcagggtggag	ccgggcaggg	aggttatggt	ggctctgggga	gtcaggggcgc	tggctcgtggg	1260	
ggactgggtg	gccaagggtgc	aggagcagct	gcagctgctg	cagggtggagc	cgggcaggga	1320	
ggttatgggtg	gtctggggag	tcagggtgct	ggctcgtggag	gccaagggtgc	aggagctgca	1380	15
gcagcagctg	cagggtggagc	cgggcaggga	ggttatggtg	gtctggggag	tcagggtgct	1440	
ggctcgtggag	gccaagggtgc	aggagctgca	gcagcagctg	cagggtggagc	cgggcaggga	1500	
ggttatgggtg	gtctgggggg	ccagggtgct	ggccaaggag	gttatgggtg	tctggggagt	1560	20
cagggcgctg	gtcgtggggg	actgggtggc	caagggtgag	gagctgctgc	tgcagctgca	1620	
gggtggagccg	ggcaggggagg	tctgggaggg	caggggagcgg	gccaagggtgc	aggagcagct	1680	
gcagcagctg	cagggtggagc	cgggcaggga	ggttatggtg	gtctggggag	tcagggtgct	1740	
ggctcgtggag	gccaagggtgc	aggagctgca	gcagcagctg	cagggtggagc	cgggcaggga	1800	25
ggttatgggtg	gtctggggag	tcagggcgct	ggctcgtggg	gactgggtgg	ccaagggtgca	1860	
ggagcagctg	cagctgctgc	aggtggagcc	ggcggaacaag	cggccgca		1908	

<210> 28

<211> 1110

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Konstrukt  
SM12

<400> 28

ggatcccagt	tacccgggca	gggaggttat	ggtaggtctg	ggggccaggg	tgctggccaa	60	45
ggaggttatg	gtggctctgg	gagtcagggc	gctggctcgt	ggggactggg	tggccaaggt	120	
gcaggagctg	ctgctgcagc	tgcagggtgga	gccgggcagg	gaggtctggg	agggcaggga	180	
gcgggccaaag	gtgcaggagc	agctgcagca	gctgcagggt	gagccgggca	gggaggttat	240	50
ggtaggtctg	ggagtcaggg	cgctggctgt	gggggactgg	gtggccaagg	tgcaggagca	300	
gctgcagctg	ctgcagggtg	agccgggcag	ggaggttatg	gtggctctgg	gagtcagggt	360	
gctggctcgt	gaggccaagg	tgcaggagct	gcagcagcag	ctgcagggtg	agccgggcag	420	
ggaggttatg	gtggctctgg	gagtcagggc	gctggctcgt	ggggactggg	tggccaaggt	480	55
gcaggagcag	ctgcagctgc	tgcagggtgga	gccgggcagg	gaggttatgg	tggctctgggg	540	
agtcagggtg	ctggctcgtg	aggccaaggt	gcaggagctg	cagcagcagc	tgcagggtgga	600	
gccgggcagg	gaggttatgg	tggctctgggg	agtcagggtg	ctggctcgtg	aggccaaggt	660	60
gcaggagctg	cagcagcagc	tgcagggtgga	gccgggcagg	gaggttatgg	tggctctgggg	720	
ggccagggtg	ctggccaagg	aggttatggt	ggctctgggga	gtcaggggcgc	tggctcgtggg	780	
ggactgggtg	gccaagggtgc	aggagctgct	gctgcagctg	cagggtggagc	cgggcaggga	840	
ggctctgggag	ggcaggggagc	gggccaaggt	gcaggagcag	ctgcagcagc	tgcagggtgga	900	65
gccgggcagg	gaggttatgg	tggctctgggg	agtcagggtg	ctggctcgtg	aggccaaggt	960	

# DE 101 13 781 A 1

gcaggagctg cagcagcagc tgcaggtgga gccgggcagg gaggttatgg tggctctggg 1020  
 agtcagggcg ctggtcgtgg gggactgggt ggccaagggtg caggagcagc tgcagctgct 1080  
 5 gcaggtggag ccggcggaca agcggccgca 1110

<210> 29  
 10 <211> 831  
 <212> DNA  
 <213> Künstliche Sequenz  
 15 <220>  
 <223> Beschreibung der künstlichen Sequenz:Konstrukt SF1

<400> 29  
 20 ggatcccagt taccgggca gggaggttat ggtggtctgg ggggccaggg tgctggccaa 60  
 ggaggttatg gtggtctggg gggccagggt gctggccaag gtgcaggagc tgctgctgca 120  
 gctgcagggtg gagccgggca gggaggttat ggtggtctgg ggagtcaggg tgctggtcgt 180  
 25 ggaggccaag gtgcaggagc tgcagcagca gctgcagggtg gagccgggca gggaggttat 240  
 ggtggtctgg ggagtcaggg cgctggtcgt gggggactgg gtggccaagg tgcaggagca 300  
 gctgcagctg ctgcagggtg agccgggcag ggaggttatg gtggtctggg gagtcagggt 360  
 gctggtcgtg gagccaagg tgcaggagct gcagcagcag ctgcagggtg agccgggcag 420  
 30 ggaggttatg gtggtctggg gagtcagggt gctggtcgtg ggggactggg tggccaagg 480  
 gcaggagcag ctgcagctgc tgcaggtgga gccgggcagg gaggttatgg tggctctggg 540  
 ggccagggtg ctggccaagg aggttatggt ggtctgggga gtcagggcgc tggctcgtgg 600  
 35 ggactgggtg gccaaagggtc aggagctgct gctgcagctg cagggtggagc cgggcaggga 660  
 ggtctgggag ggcaggagc gggccaagg gcaggagcag ctgcagcagc tgcaggtgga 720  
 gccgggcagg gaggttatgg tggctctggg agtcagggtg ctggtcgtgg aggccaagg 780  
 40 gcaggagctg cagcagcagc tgcaggtgga gccggcggac aagcgccgc a 831

<210> 30  
 45 <211> 104  
 <212> PRT  
 <213> Künstliche Sequenz  
 50 <220>  
 <223> Beschreibung der künstlichen Sequenz:SB1-Protein

<400> 30  
 55 Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly  
 1 5 10 15  
 60 Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly  
 20 25 30  
 Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln  
 65 35 40 45

# DE 101 13 781 A 1

Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala	50	55	60	
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser	65	70	75	80
Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala	85	90	95	
Ala Gly Gly Ala Gly Gln Ala Ala	100			
<210> 31				
<211> 230				
<212> PRT				
<213> Künstliche Sequenz				
<220>				
<223> Beschreibung der künstlichen Sequenz:SE1-Protein				
<400> 31				
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly	1	5	10	15
Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly	20	25	30	
Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln	35	40	45	
Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala	50	55	60	
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser	65	70	75	80
Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala	85	90	95	
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly	100	105	110	
Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala	115	120	125	
Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln	130	135	140	

# DE 101 13 781 A 1

Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala  
 145 150 155 160

5 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser  
 165 170 175

10 Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala  
 180 185 190

15 Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly  
 195 200 205

20 Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly  
 210 215 220

Gly Ala Gly Gln Ala Ala  
 225 230

25

<210> 32  
 <211> 137  
 30 <212> PRT  
 <213> Künstliche Sequenz

35 <220>  
 <223> Beschreibung der künstlichen Sequenz:SD1-Protein

<400> 32

40 Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly  
 1 5 10 15

45 Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly  
 20 25 30

Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln  
 50 35 40 45

Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala  
 50 55 60

55 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser  
 65 70 75 80

60 Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala  
 85 90 95

65 Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly

# DE 101 13 781 A 1

100	105	110	
Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala			5
115	120	125	
Ala Ala Gly Gly Ala Gly Gln Ala Ala			10
130	135		
<210> 33			
<211> 1255			15
<212> PRT			
<213> Künstliche Sequenz			
<220>			20
<223> Beschreibung der künstlichen			
Sequenz: S01S01-Protein			25
<400> 33			
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly			
1	5	10	15
			30
Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala			
20	25	30	
			35
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu			
35	40	45	
			40
Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala			
50	55	60	
			45
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser			
65	70	75	80
			50
Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala			
85	90	95	
			55
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly			
100	105	110	
			60
Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala			
115	120	125	
			65
Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln			
130	135	140	
			65
Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala			
145	150	155	160

# DE 101 13 781 A 1

Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly  
165 170 175

5 Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala  
180 185 190

10 Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala  
195 200 205

15 Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln  
210 215 220

Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly  
20 225 230 235 240

Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala  
245 250 255

25 Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly  
260 265 270

30 Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly  
275 280 285

35 Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala  
290 295 300

Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln  
40 305 310 315 320

Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala  
45 325 330 335

Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly  
340 345 350

50 Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly  
355 360 365

55 Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg  
370 375 380

60 Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly  
385 390 395 400

Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly  
65 405 410 415

# DE 101 13 781 A 1

Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala	
420 425 430	
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly	5
435 440 445	
Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln	10
450 455 460	
Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln	15
465 470 475 480	
Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly	20
485 490 495	
Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly	25
500 505 510	
Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala	30
515 520 525	
Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly	35
530 535 540	
Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly	40
545 550 555 560	
Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly	45
565 570 575	
Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala	50
580 585 590	
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly	55
595 600 605	
Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly	60
610 615 620	
Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln	65
625 630 635 640	
Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly	
645 650 655	
Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly	
660 665 670	



# DE 101 13 781 A 1

Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala  
 675 680 685  
 5 Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly  
 690 695 700  
 10 Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala  
 705 710 715 720  
 15 Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu  
 725 730 735  
 20 Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala  
 740 745 750  
 25 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser  
 755 760 765  
 30 Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala  
 770 775 780  
 35 Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly  
 785 790 795 800  
 40 Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly  
 805 810 815  
 45 Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala  
 820 825 830  
 50 Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly  
 835 840 845  
 55 Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly  
 850 855 860  
 60 Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly  
 865 870 875 880  
 65 Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr  
 885 890 895  
 70 Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu  
 900 905 910  
 75 Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly  
 915 920 925

# DE 101 13 781 A 1

Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly	
930 935 940	
Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly	5
945 950 955 960	
Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg	10
965 970 975	
Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly	15
980 985 990	
Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly	20
995 1000 1005	
Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala	25
1010 1015 1020	
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly	30
1025 1030 1035 1040	
Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly	35
1045 1050 1055	
Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg	40
1060 1065 1070	
Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly	45
1075 1080 1085	
Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly	50
1090 1095 1100	
Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly	55
1105 1110 1115 1120	
Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly	60
1125 1130 1135	
Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly	65
1140 1145 1150	
Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu	
1155 1160 1165	
Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala	
1170 1175 1180	

# DE 101 13 781 A 1

Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala  
 1185 1190 1195 1200  
 5 Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly  
 1205 1210 1215  
 10 Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg  
 1220 1225 1230  
 15 Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly  
 1235 1240 1245  
 Gly Ala Gly Gly Gln Ala Ala  
 20 1250 1255  
 <210> 34  
 25 <211> 989  
 <212> PRT  
 <213> Künstliche Sequenz  
 30 <220>  
 <223> Beschreibung der künstlichen  
 Sequenz:SO1SM12-Protein  
 35 <400> 34  
 Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly  
 1 5 10 15  
 40 Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala  
 20 25 30  
 45 Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu  
 35 40 45  
 50 Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala  
 50 55 60  
 55 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser  
 65 70 75 80  
 Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala  
 60 85 90 95  
 Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly  
 100 105 110  
 65

# DE 101 13 781 A 1

Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala	5
115 120 125	
Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln	
130 135 140	
Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala	10
145 150 155 160	
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly	15
165 170 175	
Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala	
180 185 190	20
Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala	
195 200 205	25
Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln	
210 215 220	30
Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly	
225 230 235 240	
Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala	35
245 250 255	
Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly	40
260 265 270	
Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly	45
275 280 285	
Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala	50
290 295 300	
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln	
305 310 315 320	
Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala	55
325 330 335	
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly	60
340 345 350	
Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly	65
355 360 365	

# DE 101 13 781 A 1

Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg  
370 375 380

5 Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly  
385 390 395 400

10 Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly  
405 410 415

15 Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala  
420 425 430

20 Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly  
435 440 445

25 Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln  
450 455 460

Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln  
465 470 475 480

30 Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly  
485 490 495

35 Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly  
500 505 510

40 Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala  
515 520 525

45 Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly  
530 535 540

Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly  
545 550 555 560

50 Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly  
565 570 575

55 Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala  
580 585 590

60 Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly  
595 600 605

65 Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly  
610 615 620

DE 101 13 781 A 1

Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln	625	630	635	640	
Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu		645	650	655	5
Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly		660	665	670	10
Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala		675	680	685	15
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly		690	695	700	20
Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala	705	710	715	720	25
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu		725	730	735	30
Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala		740	745	750	35
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser		755	760	765	40
Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala		770	775	780	45
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly	785	790	795	800	50
Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala		805	810	815	55
Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln		820	825	830	60
Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala		835	840	845	65
Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala		850	855	860	
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly	865	870	875	880	

# DE 101 13 781 A 1

Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly  
885 890 895

5 Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly  
900 905 910

10 Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly  
915 920 925

15 Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala  
930 935 940

20 Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly  
945 950 955 960

Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala  
965 970 975

25 Ala Ala Ala Ala Ala Gly Gly Ala Gly Gly Gln Ala Ala  
980 985

30

<210> 35  
<211> 1880  
35 <212> PRT  
<213> Künstliche Sequenz

<220>  
40 <223> Beschreibung der künstlichen  
Sequenz:SO1SO1SO1-Protein

45 <400> 35  
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly  
1 5 10 15

50 Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala  
20 25 30

55 Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu  
35 40 45

Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala  
60 50 55 60

65 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser  
65 70 75 80

Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala

## DE 101 13 781 A 1

	85	90	95	
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly	100	105	110	5
Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala	115	120	125	10
Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln	130	135	140	15
Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala	145	150	155	20
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly	165	170	175	25
Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala	180	185	190	30
Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala	195	200	205	35
Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln	210	215	220	40
Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly	225	230	235	45
Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala	245	250	255	50
Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly	260	265	270	55
Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly	275	280	285	60
Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala	290	295	300	65
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln	305	310	315	
Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala	325	330	335	
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly				



# DE 101 13 781 A 1

	340	345	350
5	Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly 355	360	365
10	Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg 370	375	380
15	Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly 385	390	395 400
20	Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly 405	410	415
25	Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala 420	425	430
30	Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly 435	440	445
35	Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln 450	455	460
40	Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln 465	470	475 480
45	Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly 485	490	495
50	Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly 500	505	510
55	Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala 515	520	525
60	Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly 530	535	540
65	Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly 545	550	555 560
	Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly 565	570	575
	Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala 580	585	590
	Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly		

# DE 101 13 781 A 1

595	600	605	
Gly Leu Gly Gly Gln Gly	Ala Gly Ala Ala Ala Ala Ala	Gly Gly	5
610	615	620	
Ala Gly Gln Gly Gly Tyr Gly	Gly Gly Leu Gly Gly Gln Gly	Ala Gly Gln	10
625	630	635 640	
Gly Gly Tyr Gly Gly Leu Gly	Gly Gly Gln Gly Ala Gly	Gln Gly Ala Gly	15
	645	650 655	
Ala Ala Ala Ala Ala Ala	Gly Gly Ala Gly Gln Gly	Gly Tyr Gly Gly	20
	660	665 670	
Leu Gly Ser Gln Gly Ala Gly	Arg Gly Gly Gln Gly Ala	Gly Ala Ala	25
	675	680 685	
Ala Ala Ala Ala Gly Gly	Ala Gly Gln Gly Gly Tyr	Gly Gly Leu Gly	30
	690	695 700	
Ser Gln Gly Ala Gly Arg Gly	Gly Gly Leu Gly Gly Gln Gly	Ala Gly Ala	35
705	710	715 720	
Ala Ala Ala Ala Ala Gly	Gly Gly Ala Gly Gln Gly	Gly Tyr Gly Gly Leu	40
	725	730 735	
Gly Ser Gln Gly Ala Gly	Arg Gly Gly Gln Gly Ala	Gly Ala Ala Ala	45
	740	745 750	
Ala Ala Ala Gly Gly Ala Gly	Gln Gly Gly Tyr Gly Gly	Leu Gly Ser	50
	755	760 765	
Gln Gly Ala Gly Arg Gly	Gly Gly Leu Gly Gly Gln Gly	Ala Gly Ala Ala	55
	770	775 780	
Ala Ala Ala Ala Gly Gly	Ala Gly Gln Gly Gly Tyr	Gly Gly Leu Gly	60
785	790	795 800	
Gly Gln Gly Ala Gly Gln	Gly Gly Tyr Gly Gly Leu	Gly Ser Gln Gly	65
	805	810 815	
Ala Gly Arg Gly Gly Leu	Gly Gly Gln Gly Ala Gly	Ala Ala Ala Ala	70
	820	825 830	
Ala Ala Gly Gly Ala Gly	Gln Gly Gly Leu Gly Gly	Gln Gly Ala Gly	75
	835	840 845	
Gln Gly Ala Gly Ala Ala	Ala Ala Ala Gly Gly	Ala Gly Gln Gly	80

## DE 101 13 781 A 1

	850	855	860
5	Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly		
	865	870	875 880
10	Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr		
	885	890	895
15	Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu		
	900	905	910
20	Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly		
	915	920	925
25	Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly		
	930	935	940
30	Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly		
	945	950	955 960
35	Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg		
	965	970	975
40	Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly		
	980	985	990
45	Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly		
	995	1000	1005
50	Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala		
	1010	1015	1020
55	Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly		
	1025	1030	1035 1040
60	Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly		
	1045	1050	1055
65	Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg		
	1060	1065	1070
70	Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly		
	1075	1080	1085
75	Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly		
	1090	1095	1100
80	Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly		

# DE 101 13 781 A 1

1105	1110	1115	1120	
Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly				5
1125	1130	1135		
Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly				10
1140	1145	1150		
Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu				15
1155	1160	1165		
Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala				20
1170	1175	1180		
Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala				25
1185	1190	1195	1200	
Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly				30
1205	1210	1215		
Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg				35
1220	1225	1230		
Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly				40
1235	1240	1245		
Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly				45
1250	1255	1260		
Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala				50
1265	1270	1275	1280	
Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly				55
1285	1290	1295		
Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala				60
1300	1305	1310		
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu				65
1315	1320	1325		
Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly				
1330	1335	1340		
Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly				
1345	1350	1355	1360	
Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala				

# DE 101 13 781 A 1

	1365	1370	1375
5	Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly 1380	1385	1390
10	Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala 1395	1400	1405
15	Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu 1410	1415	1420
20	Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln 1425	1430	1435 1440
25	Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala 1445	1450	1455
30	Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala 1460	1465	1470
35	Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln 1475	1480	1485
40	Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln 1490	1495	1500
45	Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly 1505	1510	1515 1520
50	Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly 1525	1530	1535
55	Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala 1540	1545	1550
60	Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly 1555	1560	1565
65	Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly 1570	1575	1580
	Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly 1585	1590	1595 1600
	Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala 1605	1610	1615
	Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala		

# DE 101 13 781 A 1

1620	1625	1630	
Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly			5
1635	1640	1645	
Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg			10
1650	1655	1660	
Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly			15
1665	1670	1675	1680
Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly			
1685	1690	1695	
Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala			20
1700	1705	1710	
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly			25
1715	1720	1725	
Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln			30
1730	1735	1740	
Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr			35
1745	1750	1755	1760
Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln			40
1765	1770	1775	
Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly			
1780	1785	1790	
Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala			45
1795	1800	1805	
Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly			50
1810	1815	1820	
Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly			55
1825	1830	1835	1840
Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly			60
1845	1850	1855	
Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala			
1860	1865	1870	
Gly Gly Ala Gly Gly Gln Ala Ala			65

# DE 101 13 781 A 1

1875

1880

5

<210> 36

<211> 219

<212> PRT

10 <213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:FA2-Protein

15

<400> 36

Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly

20

1

5

10

15

Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly

20

25

30

25

Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln

35

40

45

30

Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala

50

55

60

35

Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser

65

70

75

80

Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala

40

85

90

95

Ala Ala Ala Ala Gly Gly Ala Gly Ser Gly Ser Gly Ala Gly Ala Gly

45

100

105

110

Ser Gly Ala Gly Ala Gly Ala Gly Ser Gly Ala Gly Ala Gly Ser Gly

115

120

125

50

Ala Gly Ala Gly Ser Gly Ala Gly Ala Gly Ser Gly Ala Gly Ala Gly

130

135

140

55

Tyr Gly Ala Gly Ala Gly Val Gly Tyr Gly Ala Gly Tyr Gly Ala Gly

145

150

155

160

60

Ala Gly Val Gly Tyr Gly Ala Gly Ala Gly Ser Gly Ala Ala Ser Gly

165

170

175

Ala Gly Ala Gly Ala Gly Ala Gly Thr Gly Ser Ser Gly Phe Gly Pro

65

180

185

190

# DE 101 13 781 A 1

Tyr Val Ala Asn Gly Gly Tyr Ser Gly Tyr Glu Tyr Ala Trp Ser Ser  
 195 200 205

Lys Ser Asp Phe Glu Thr Ala Gly Gln Ala Ala  
 210 215

<210> 37

<211> 170

<212> PRT

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:SA1-Protein

<400> 37

Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly  
 1 5 10 15

Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala  
 20 25 30

Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu  
 35 40 45

Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala  
 50 55 60

Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser  
 65 70 75 80

Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala  
 85 90 95

Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly  
 100 105 110

Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala  
 115 120 125

Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln  
 130 135 140

Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala  
 145 150 155 160

Ala Ala Ala Gly Gly Ala Gly Gln Ala Ala  
 165 170



# DE 101 13 781 A 1

<210> 38  
 <211> 630  
 <212> PRT  
 5 <213> Künstliche Sequenz  
  
 <220>  
 10 <223> Beschreibung der künstlichen Sequenz:SO1-Protein  
  
 <400> 38  
 Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly  
 15 1 5 10 15  
  
 Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala  
 20 20 25 30  
  
 Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu  
 25 35 40 45  
  
 Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala  
 30 50 55 60  
  
 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser  
 35 65 70 75 80  
  
 Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala  
 40 85 90 95  
  
 Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly  
 45 100 105 110  
  
 Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala  
 50 115 120 125  
  
 Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln  
 55 130 135 140  
  
 Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala  
 60 145 150 155 160  
  
 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly  
 65 165 170 175  
  
 Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala  
 180 185 190  
  
 Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala

# DE 101 13 781 A 1

195	200	205	
Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln			5
210	215	220	
Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly			10
225	230	235	240
Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala			15
245	250	255	
Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly			20
260	265	270	
Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly			25
275	280	285	
Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala			30
290	295	300	
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln			35
305	310	315	320
Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala			40
325	330	335	
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly			45
340	345	350	
Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly			50
355	360	365	
Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg			55
370	375	380	
Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly			60
385	390	395	400
Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly			65
405	410	415	
Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala			
420	425	430	
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly			
435	440	445	
Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln			

# DE 101 13 781 A 1

450                      455                      460  
 5 Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln  
 465                      470                      475                      480  
 Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly  
 10                      485                      490                      495  
 Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly  
 15                      500                      505                      510  
 Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala  
 20                      515                      520                      525  
 Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly  
 25                      530                      535                      540  
 Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly  
 30                      545                      550                      555                      560  
 Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly  
 35                      565                      570                      575  
 Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala  
 40                      580                      585                      590  
 Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly  
 45                      595                      600                      605  
 Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly  
 50                      610                      615                      620  
 Ala Gly Gly Gln Ala Ala  
 55                      625                      630  
 <210> 39  
 <211> 364  
 <212> PRT  
 55 <213> Künstliche Sequenz  
 <220>  
 60 <223> Beschreibung der künstlichen Sequenz: SM12-Protein  
 <400> 39  
 Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly  
 65                      1                      5                      10                      15

# DE 101 13 781 A 1

Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly	
20 25 30	
Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln	5
35 40 45	
Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala	10
50 55 60	
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser	15
65 70 75 80	
Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala	20
85 90 95	
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly	25
100 105 110	
Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala	30
115 120 125	
Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln	35
130 135 140	
Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala	40
145 150 155 160	
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser	45
165 170 175	
Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala	50
180 185 190	
Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly	55
195 200 205	
Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly	60
210 215 220	
Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly	65
225 230 235 240	
Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly	
245 250 255	
Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala	
260 265 270	

# DE 101 13 781 A 1

Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala  
275 280 285

5 Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu  
290 295 300

10 Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala  
305 310 315 320

15 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser  
325 330 335

20 Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala  
340 345 350

Ala Ala Ala Ala Gly Gly Ala Gly Gly Gln Ala Ala  
355 360

25

<210> 40  
<211> 271  
30 <212> PRT  
<213> Künstliche Sequenz

35 <220>  
<223> Beschreibung der künstlichen Sequenz:SF1-Protein

<400> 40  
40 Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly  
1 5 10 15

45 Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala  
20 25 30

Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu  
50 35 40 45

Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala  
50 55 60

55 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser  
65 70 75 80

60 Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala  
85 90 95

65 Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly  
100 105 110

# DE 101 13 781 A 1

Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala	5
115 120 125	
Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln	
130 135 140	
Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala	10
145 150 155 160	
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly	15
165 170 175	
Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala	
180 185 190	20
Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala	
195 200 205	25
Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln	
210 215 220	30
Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly	
225 230 235 240	
Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala	35
245 250 255	
Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gly Gln Ala Ala	40
260 265 270	
<210> 41	45
<211> 182	
<212> DNA	
<213> Künstliche Sequenz	50
<220>	
<223> Beschreibung der künstlichen Sequenz:ELP mit 10 Pentamereinheiten	55
<400> 41	
ctcgagatgg gccacggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg 60	
ggcgccaggtg ttctgtgtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca 120	60
gggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccggggcg gctggcgggcc 180	
gc 182	65

<210> 42  
 <211> 332  
 <212> DNA  
 5 <213> Künstliche Sequenz

<220>  
 10 <223> Beschreibung der künstlichen Sequenz:ELP mit 20  
 Pentamereinheiten

<400> 42  
 15 ctcgagatgg gccacggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg 60  
 ggcgcaggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca 120  
 ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg 180  
 20 ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcaggtg ttccctggtgt aggtgtgccg 240  
 ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg 300  
 ggtggcgggtg tgccgggcgg gctggcggcc gc 332

25  
 <210> 43  
 <211> 482  
 <212> DNA  
 30 <213> Künstliche Sequenz

<220>  
 35 <223> Beschreibung der künstlichen Sequenz:ELP mit 30  
 Pentamereinheiten

<400> 43  
 40 ctcgagatgg gccacggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg 60  
 ggcgcaggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca 120  
 ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg 180  
 45 ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcaggtg ttccctggtgt aggtgtgccg 240  
 ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg 300  
 ggtggcgggtg tgccgggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg 360  
 ggcgcaggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca 420  
 50 ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgg gctggcggcc 480  
 gc 482

55 <210> 44  
 <211> 632  
 <212> DNA  
 60 <213> Künstliche Sequenz

<220>  
 <223> Beschreibung der künstlichen Sequenz:ELP mit 40  
 65 Pentamereinheiten

# DE 101 13 781 A 1

<400> 44

ctcagatgg gccacggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg	60	
ggcgcaggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca	120	5
ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg	180	
ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcaggtg ttccctggtgt aggtgtgccg	240	
ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg	300	
ggtggcgggtg tgccgggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg	360	10
ggcgcaggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca	420	
ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg	480	
ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcaggtg ttccctggtgt aggtgtgccg	540	
ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg	600	15
ggtggcgggtg tgccgggcgg gctggcggcc gc	632	

<210> 45

<211> 932

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:ELP mit 60  
Pentamereinheiten

<400> 45

ctcagatgg gccacggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg	60	35
ggcgcaggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca	120	
ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg	180	
ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcaggtg ttccctggtgt aggtgtgccg	240	
ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg	300	40
ggtggcgggtg tgccgggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg	360	
ggcgcaggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca	420	
ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg	480	45
ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcaggtg ttccctggtgt aggtgtgccg	540	
ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg	600	
ggtggcgggtg tgccgggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg	660	
ggcgcaggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca	720	50
ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg	780	
ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcaggtg ttccctggtgt aggtgtgccg	840	
ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg	900	55
ggtggcgggtg tgccgggcgg gctggcggcc gc	932	

<210> 46

<211> 1082

<212> DNA

<213> Künstliche Sequenz

<220>



# DE 101 13 781 A 1

<223> Beschreibung der künstlichen Sequenz:ELP mit 70  
Pentamereinheiten

5 <400> 46  
ctcagagatgg gccacggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg 60  
ggcgcagggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca 120  
10 ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg 180  
ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcagggtg ttccctggtgt aggtgtgccg 240  
ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg 300  
ggtggcgggtg tgccgggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg 360  
15 ggcgcagggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca 420  
ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg 480  
ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcagggtg ttccctggtgt aggtgtgccg 540  
20 ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg 600  
ggtggcgggtg tgccgggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg 660  
ggcgcagggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca 720  
ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg 780  
25 ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcagggtg ttccctggtgt aggtgtgccg 840  
ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg 900  
ggtggcgggtg tgccgggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg 960  
30 ggcgcagggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca 1020  
ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggtggcgggc 1080  
gc 1082

35

<210> 47

<211> 1532

<212> DNA

40

<213> Künstliche Sequenz

<220>

45 <223> Beschreibung der künstlichen Sequenz:ELP mit 100  
Pentamereinheiten

<400> 47

50 ctcagagatgg gccacggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg 60  
ggcgcagggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca 120  
ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg 180  
55 ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcagggtg ttccctggtgt aggtgtgccg 240  
ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg 300  
ggtggcgggtg tgccgggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg 360  
ggcgcagggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca 420  
60 ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg 480  
ggcgtgggtg ttccgggtgg cgggtgtgccg ggcgcagggtg ttccctggtgt aggtgtgccg 540  
ggtgttggtg tgccgggtgt tgggtgtacca ggtggcgggtg ttccgggtgc aggcgttccg 600  
ggtggcgggtg tgccgggcgt ggggtgttccg ggcgtgggtg ttccgggtgg cgggtgtgccg 660  
65 ggcgcagggtg ttccctggtgt aggtgtgccg ggtgttggtg tgccgggtgt tgggtgtacca 720  
ggtggcgggtg ttccgggtgc aggcgttccg ggtggcgggtg tgccgggcgt ggggtgttccg 780

# DE 101 13 781 A 1

ggcgtgggtg	ttccgggtgg	cggtgtgccg	ggcgcaggtg	ttcctggtgt	aggtgtgccg	840	
ggtgttggtg	tgccgggtgt	tggtgtacca	ggtggcgggtg	ttccgggtgc	aggcgttccg	900	
ggtggcgggtg	tgccgggcgt	gggtgttccg	ggcgtgggtg	ttccgggtgg	cggtgtgccg	960	5
ggcgcaggtg	ttcctggtgt	aggtgtgccg	ggtgttggtg	tgccgggtgt	tggtgtacca	1020	
ggtggcgggtg	ttccgggtgc	aggcgttccg	ggtggcgggtg	tgccgggcgt	gggtgttccg	1080	
ggcgtgggtg	ttccgggtgg	cggtgtgccg	ggcgcaggtg	ttcctggtgt	aggtgtgccg	1140	
ggtgttggtg	tgccgggtgt	tggtgtacca	ggtggcgggtg	ttccgggtgc	aggcgttccg	1200	10
ggtggcgggtg	tgccgggcgt	gggtgttccg	ggcgtgggtg	ttccgggtgg	cggtgtgccg	1260	
ggcgcaggtg	ttcctggtgt	aggtgtgccg	ggtgttggtg	tgccgggtgt	tggtgtacca	1320	
ggtggcgggtg	ttccgggtgc	aggcgttccg	ggtggcgggtg	tgccgggcgt	gggtgttccg	1380	
ggcgtgggtg	ttccgggtgg	cggtgtgccg	ggcgcaggtg	ttcctggtgt	aggtgtgccg	1440	15
ggtgttggtg	tgccgggtgt	tggtgtacca	ggtggcgggtg	ttccgggtgc	aggcgttccg	1500	
ggtggcgggtg	tgccgggcgt	gctggcgcc	gc			1532	

<210> 48

<211> 2322

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: SM12-70xELP  
(Pflanzen)

<400> 48

atggtttcca	aaccttttct	atctttgctt	tacttttctt	tgcttctctt	tacaagcaca	60	
tggttagcag	gatcccagtt	acccgggcag	ggaggttatg	gtggtctggg	gggccagggt	120	
gctggccaag	gaggttatgg	tggtctgggg	agtcagggcg	ctggtcgtgg	gggactgggt	180	
ggccaagggtg	caggagctgc	tgctgcagct	gcaggtggag	ccgggcaggg	aggtctggga	240	40
gggcaggggag	cgggccaagg	tgaggagca	gctgcagcag	ctgcaggtgg	agccgggcag	300	
ggaggttatg	gtggtctggg	gagtcagggc	gctggctgtg	ggggactggg	tgccaagggt	360	
gcaggagcag	ctgcagctgc	tgagggtgga	gccgggcagg	gaggttatgg	tggtctgggg	420	45
agtcagggtg	ctggtcgtgg	aggccaagggt	gcaggagctg	cagcagcagc	tgagggtgga	480	
gccgggcagg	gaggttatgg	tggtctgggg	agtcagggcg	ctggtcgtgg	gggactgggt	540	
ggccaagggtg	caggagcagc	tgagctgct	gcaggtggag	ccgggcaggg	aggttatggt	600	
ggtctggggga	gtcagggtgc	tggtcgtgga	ggccaagggtg	caggagctgc	agcagcagct	660	50
gcagggtggag	ccgggcaggg	aggttatggt	ggtctggggga	gtcagggtgc	tggtcgtgga	720	
ggccaagggtg	caggagctgc	agcagcagct	gcaggtggag	ccgggcaggg	aggttatggt	780	
ggtctggggg	gccagggtgc	tgccaaggga	ggttatggtg	gtctggggag	tcagggcgct	840	55
ggtcgtgggg	gactgggtgg	ccaagggtgca	ggagctgctg	ctgcagctgc	aggtggagcc	900	
gggcaggggag	gtctgggagg	gcagggagcg	ggccaagggtg	caggagcagc	tgagcagct	960	
gcagggtggag	ccgggcaggg	aggttatggt	ggtctggggga	gtcagggtgc	tggtcgtgga	1020	
ggccaagggtg	caggagctgc	agcagcagct	gcaggtggag	ccgggcaggg	aggttatggt	1080	60
ggtctggggga	gtcagggcgc	tggtcgtggg	ggactgggtg	gccaagggtgc	aggagcagct	1140	
gcagctgctg	cagggtggagc	cggcggaaca	gcggccgcag	aacaaaaact	catctcagaa	1200	
gaggatctga	atggggccgt	cgagatgggc	cacggcgtgg	gtgttccggg	cgtgggtgtt	1260	
ccgggtggcg	gtgtgccggg	cgcaggtgtt	cctgggtgtg	gtgtgccggg	tggtggtgtg	1320	65
ccgggtgttg	gtgtaccagg	tggggtgtt	ccgggtgcag	gcgttccggg	tggggtgtg	1380	

# DE 101 13 781 A 1

```

ccgggcgtgg gtgttccggg cgtgggtggt ccgggtggcg gtgtgccggg cgcaggtggt 1440
cctggtgtag gtgtgccggg tgttggtgtg ccgggtggtg gtgtaccagg tggcgggtgtt 1500
ccgggtgcag gcgttccggg tggcgggtgtg ccgggcgtgg gtgttccggg cgtgggtggt 1560
5 ccgggtggcg gtgtgccggg cgcaggtggt cctggtgtag gtgtgccggg tgttggtgtg 1620
ccgggtggtg gtgtaccagg tggcgggtgtt ccgggtgcag gcgttccggg tggcgggtgtg 1680
ccgggcgtgg gtgttccggg cgtgggtggt ccgggtggcg gtgtgccggg cgcaggtggt 1740
10 cctggtgtag gtgtgccggg tgttggtgtg ccgggtggtg gtgtaccagg tggcgggtgtt 1800
ccgggtgcag gcgttccggg tggcgggtgtg ccgggcgtgg gtgttccggg cgtgggtggt 1860
ccgggtggcg gtgtgccggg cgcaggtggt cctggtgtag gtgtgccggg tgttggtgtg 1920
ccgggtggtg gtgtaccagg tggcgggtgtt ccgggtgcag gcgttccggg tggcgggtgtg 1980
15 ccgggcgtgg gtgttccggg cgtgggtggt ccgggtggcg gtgtgccggg cgcaggtggt 2040
cctggtgtag gtgtgccggg tgttggtgtg ccgggtggtg gtgtaccagg tggcgggtgtt 2100
ccgggtgcag gcgttccggg tggcgggtgtg ccgggcgtgg gtgttccggg cgtgggtggt 2160
20 ccgggtggcg gtgtgccggg cgcaggtggt cctggtgtag gtgtgccggg tgttggtgtg 2220
ccgggtggtg gtgtaccagg tggcgggtgtt ccgggtgcag gcgttccggg tggcgggtgtg 2280
ccgggcgggc tggcggccgc agaaccctct ag 2322

```

25

<210> 49

<211> 773

<212> PRT

30

<213> Künstliche Sequenz

<220>

35 <223> Beschreibung der künstlichen Sequenz: SM12-70xELP  
(Pflanzen)

<400> 49

40

Met Ala Ser Lys Pro Phe Leu Ser Leu Leu Ser Leu Ser Leu Leu Leu

1

5

10

15

45

Phe Thr Ser Thr Cys Leu Ala Gly Ser Gln Leu Pro Gly Gln Gly Gly

20

25

30

50

Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly

35

40

45

55

Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala

50

55

60

60

Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly

65

70

75

80

65

Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly

85

90

95

65

Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly

100

105

110

# DE 101 13 781 A 1

Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala	115	120	125	
Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala	130	135	140	5
Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly	145	150	155	10
Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg	165	170	175	15
Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly	180	185	190	20
Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly	195	200	205	25
Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala	210	215	220	30
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly	225	230	235	35
Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln	245	250	255	40
Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr	260	265	270	45
Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln	275	280	285	50
Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly	290	295	300	55
Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala	305	310	315	60
Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly	325	330	335	65
Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly	340	345	350	
Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly	355	360	365	

# DE 101 13 781 A 1

Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala  
 370 375 380

5 Gly Gly Ala Gly Gly Gln Ala Ala Ala Glu Gln Lys Leu Ile Ser Glu  
 385 390 395 400

10 Glu Asp Leu Asn Gly Ala Val Glu Met Gly His Gly Val Gly Val Pro  
 405 410 415

15 Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly  
 420 425 430

20 Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly  
 435 440 445

25 Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly  
 450 455 460

30 Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val  
 465 470 475 480

35 Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro  
 485 490 495

40 Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly  
 500 505 510

45 Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala  
 515 520 525

50 Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly  
 530 535 540

55 Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val  
 545 550 555 560

60 Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro  
 565 570 575

65 Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly  
 580 585 590

60 Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly  
 595 600 605

65 Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly  
 610 615 620

DE 101 13 781 A 1

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val	
625                          630                          635                          640	
Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro	5
645                          650                          655	
Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly	10
660                          665                          670	
Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val	15
675                          680                          685	
Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly	20
690                          695                          700	
Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val	25
705                          710                          715                          720	
Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro	30
725                          730                          735	
Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly	35
740                          745                          750	
Ala Gly Val Pro Gly Gly Gly Val Pro Gly Gly Leu Ala Ala Ala Glu	40
755                          760                          765	
Pro Lys Asp Glu Leu	45
770	
<210> 50	45
<211> 2334	
<212> DNA	
<213> Künstliche Sequenz	50
<220>	
<223> Beschreibung der künstlichen Sequenz:SM12-70xELP (E.coli)	55
<400> 50	
atggctagca tgactggttg acagcaaattg ggtcgccgat cccagttagc cgaggcaggga 60	
ggttatggtg gtctgggggg ccagggtgct ggccaaggag gttatggtg tctggggagt 120	60
cagggcgctg gtcgtggggg actgggtggc caagggtgcag gagctgctgc tgcagctgca 180	
ggtggagccg gccaggggagg tctggggagg caggggagcgg gccaaagggtgc aggagcagct 240	
gcagcagctg caggtgggagc cgggcaggga ggttatggtg gtctggggag tcagggcgct 300	
ggtcgtgggg gactgggttg ccaagggtgca ggagcagctg agtctgctgc aggtggagcc 360	65

# DE 101 13 781 A 1

```

gggcagggag gttatggtgg tctggggagt caggggtgctg gtcgtggagg ccaaggtgca 420
ggagctgcag cagcagctgc aggtggagcc gggcagggag gttatggtgg tctggggagt 480
cagggcgctg gtcgtggggg actgggtggc caaggtgcag gagcagctgc agctgctgca 540
5 ggtggagccg ggcagggagg ttatggtggt ctggggagtc aggggtgctgg tcgtggaggc 600
caaggtgcag gagctgcagc agcagctgca ggtggagccg ggcagggagg ttatggtggt 660
ctggggagtc aggggtgctgg tcgtggaggc caaggtgcag gagctgcagc agcagctgca 720
10 ggtggagccg ggcagggagg ttatggtggt ctggggggcc aggggtgctgg ccaaggaggt 780
tatggtggtc tggggagtca gggcgctggt cgtggggggac tgggtggcca aggtgcagga 840
gctgctgctg cagctgcagg tggagccggg cagggaggtc tgggagggca gggagcgggc 900
caaggtgcag gagcagctgc agcagctgca ggtggagccg ggcagggagg ttatggtggt 960
15 ctggggagtc aggggtgctgg tcgtggaggc caaggtgcag gagctgcagc agcagctgca 1020
ggtggagccg ggcagggagg ttatggtggt ctggggagtc agggcgctgg tcgtggggga 1080
ctgggtggcc aaggtgcagg agcagctgca gctgctgag gtggagccgg cggacaagcg 1140
20 gccgcagAAC aaaaactcat ctCagaagag gatctgaatg gggccgtcga gatgggccac 1200
ggcgtgggtg ttccgggcgt ggggtgttccg ggtggcggtg tgccggggcg aggtgttcc 1260
ggtgtaggtg tgccgggtgt tgggtgtgccg ggtgttgggtg taccaggtgg cgggtgttcc 1320
ggtgcaggcg ttccgggtgg cgggtgtgccg ggcgtgggtg ttccgggcgt ggggtgttcc 1380
25 ggtggcggtg tgccggggcg aggtgttcc  ggtgtaggtg tgccgggtgt tgggtgtgcc 1440
ggtgttgggtg taccaggtgg cgggtgttcc ggtgcaggcg ttccgggtgg cgggtgtgcc 1500
ggcgtgggtg ttccgggcgt ggggtgttcc ggtggcggtg tgccggggcg aggtgttcc 1560
ggtgtaggtg tgccgggtgt tgggtgtgcc ggtgttgggtg taccaggtgg cgggtgttcc 1620
30 ggtgcaggcg ttccgggtgg cgggtgtgcc ggcgtgggtg ttccgggcgt ggggtgttcc 1680
ggtggcggtg tgccggggcg aggtgttcc ggtgtaggtg tgccgggtgt tgggtgtgcc 1740
ggtgttgggtg taccaggtgg cgggtgttcc ggtgcaggcg ttccgggtgg cgggtgtgcc 1800
35 ggcgtgggtg ttccgggcgt ggggtgttcc ggtggcggtg tgccggggcg aggtgttcc 1860
ggtgtaggtg tgccgggtgt tgggtgtgcc ggtgttgggtg taccaggtgg cgggtgttcc 1920
ggtgcaggcg ttccgggtgg cgggtgtgcc ggcgtgggtg ttccgggcgt ggggtgttcc 1980
ggtggcggtg tgccggggcg aggtgttcc ggtgtaggtg tgccgggtgt tgggtgtgcc 2040
40 ggtgttgggtg taccaggtgg cgggtgttcc ggtgcaggcg ttccgggtgg cgggtgtgcc 2100
ggcgtgggtg ttccgggcgt ggggtgttcc ggtggcggtg tgccggggcg aggtgttcc 2160
ggtgtaggtg tgccgggtgt tgggtgtgcc ggtgttgggtg taccaggtgg cgggtgttcc 2220
45 ggtgcaggcg ttccgggtgg cgggtgtgcc ggcgggctgg cggccgcaga acaaaaactc 2280
atctcagaag aggatctgaa tggggccgtc gagcaccacc accaccacca ctga 2334

```

50 <210> 51

<211> 777

<212> PRT

<213> Künstliche Sequenz

55

<220>

<223> Beschreibung der künstlichen Sequenz: SM12-70xELP

60

(E.coli)

<400> 51

Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Gln Leu

65

1

5

10

15

# DE 101 13 781 A 1

Pro Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln	20	25	30	
Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu	35	40	45	5
Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly	50	55	60	10
Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala	65	70	75	80
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly	85	90	95	20
Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala	100	105	110	25
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu	115	120	125	30
Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala	130	135	140	35
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser	145	150	155	40
Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala	165	170	175	45
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly	180	185	190	50
Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala	195	200	205	55
Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln	210	215	220	60
Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala	225	230	235	65
Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala	245	250	255	
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly	260	265	270	



# DE 101 13 781 A 1

Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly  
 275 280 285  
 5 Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly  
 290 295 300  
 10 Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly  
 305 310 315 320  
 15 Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala  
 325 330 335  
 20 Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly  
 340 345 350  
 25 Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala  
 355 360 365  
 30 Ala Ala Ala Ala Ala Gly Gly Ala Gly Gly Gln Ala Ala Ala Glu Gln  
 370 375 380  
 35 Lys Leu Ile Ser Glu Glu Asp Leu Asn Gly Ala Val Glu Met Gly His  
 385 390 395 400  
 40 Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly  
 405 410 415  
 45 Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val  
 420 425 430  
 50 Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly  
 435 440 445  
 55 Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val  
 450 455 460  
 60 Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro  
 465 470 475 480  
 65 Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly  
 485 490 495  
 70 Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly  
 500 505 510  
 75 Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly  
 515 520 525

# DE 101 13 781 A 1

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val	530	535	540	
Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro	545	550	555	560
Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly	565	570	575	
Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala	580	585	590	
Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly	595	600	605	
Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val	610	615	620	
Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro	625	630	635	640
Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly	645	650	655	
Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val	660	665	670	
Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly	675	680	685	
Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val	690	695	700	
Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro	705	710	715	720
Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly	725	730	735	
Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Gly	740	745	750	
Leu Ala Ala Ala Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Gly	755	760	765	
Ala Val Glu His His His His His His	770	775		

- 66

sequenzen.

23. Verfahren zur Herstellung von Spinnenseidenprotein-produzierenden Pflanzen bzw. Pflanzenzellen, umfassend die folgenden Schritte:
  - a) Herstellung eines rekombinanten Nukleinsäuremoleküls nach einem der Ansprüche 7 bis 17;
  - b) Übertragung des Nukleinsäuremoleküls aus a) auf pflanzliche Zellen; und
  - c) ggf. die Regeneration fertiler Pflanzen aus den transformierten Pflanzenzellen.
24. Transgene Pflanzenzellen, enthaltend ein rekombinantes Nukleinsäuremolekül oder einen Vektor nach einem der Ansprüche 7 bis 18, oder hergestellt nach einem Verfahren nach Anspruch 23.
25. Transgene Pflanzen, enthaltend eine Pflanzenzelle nach Anspruch 24 oder hergestellt nach Anspruch 23, sowie Teile dieser Pflanzen, transgene Ernteprodukte und transgenes Vermehrungsmaterial dieser Pflanzen, wie Protoplasten, Pflanzenzellen, Kalli, Samen, Knollen, Stecklinge, sowie die transgenen Nachkommen dieser Pflanzen.
26. Transgene Pflanzen nach Anspruch 25, ausgewählt aus der Gruppe, bestehend aus Tabakpflanze und Kartoffelpflanze.
27. Verfahren zur Gewinnung von pflanzlichem Spinnenseidenprotein, umfassend die folgenden Schritte:
  - a) die Übertragung eines rekombinanten Nukleinsäuremoleküls oder Vektors nach einem der Ansprüche 7 bis 18 auf Pflanzenzellen;
  - b) gegebenenfalls die Regeneration von Pflanzen aus den transformierten Pflanzenzellen; und
  - c) die Verarbeitung der Pflanzenzellen aus a) bzw. der Pflanzen aus b) zur Gewinnung von pflanzlichem Spinnenseidenprotein.
28. Verfahren zur Gewinnung von rekombinant hergestelltem Spinnenseidenprotein, umfassend die folgenden Schritte:
  - a) die Übertragung eines rekombinanten Nukleinsäuremoleküls oder Vektors nach einem der Ansprüche 7 bis 18 auf Zellen;
  - b) die Reinigung des Spinnenseidenproteins durch Hitzebehandlung des Zellextrakts und anschließende Abtrennung der denaturierten zelleigenen Proteine.
29. Verfahren zur Gewinnung von rekombinant hergestelltem Spinnenseidenprotein, umfassend die folgenden Schritte:
  - a) die Übertragung eines rekombinanten Nukleinsäuremoleküls oder Vektors nach einem der Ansprüche 7 bis 18 auf Zellen;
  - b) die Reinigung des pflanzlichen Spinnenseidenproteins durch Einstellung eines sauren pH, vorzugsweise eines pH im Bereich von 2,5 bis 3,5 mittels Zugabe von Säure, vorzugsweise Salzsäure zu dem Zellextrakt und anschließende Abtrennung der denaturierten zelleigenen Proteine.
30. Verfahren zur Gewinnung von rekombinant hergestelltem Spinnenseidenprotein, umfassend die folgenden Schritte:
  - a) die Übertragung eines rekombinanten Nukleinsäuremoleküls nach einem der Ansprüche 15 bis 17 auf Zellen;
  - b) die Reinigung des Spinnenseidenproteins auf folgende Weise:
    - Anreicherung des Spinnenseiden-ELP-Fusionsproteins durch Hitzebehandlung des Zellextrakts;
    - Ausfällung des Spinnenseiden-ELP-Fusionsproteins durch weitere Temperaturerhöhung, vorzugsweise auf eine Temperatur von mindestens 60°C, und vorzugsweise bei einer Salzkonzentration von 1 M bis 2 M; und
    - Abspaltung des ELP-Fragments, vorzugsweise durch Verdauung mit CNBr.
31. Verfahren nach einem der Ansprüche 28 bis 30, dadurch gekennzeichnet, daß die Zellen ausgewählt werden aus Pflanzenzellen, tierischen Zellen und bakteriellen Zellen.
32. Pflanzliches Spinnenseidenprotein, hergestellt nach einem Verfahren nach einem der Ansprüche 27 bis 31.
33. Spinnenseidenprotein nach Anspruch 32, dadurch gekennzeichnet, daß es ein Molekulargewicht im Bereich von 10 bis 160 kDa aufweist.
34. Verwendung der Spinnenseidenproteine nach einem der Ansprüche 20 bis 22 bzw. nach Anspruch 32 oder 33 zur Herstellung von synthetische Fäden, Folien und/oder Membranen.
35. Verwendung nach Anspruch 34, wobei die Fäden, Folien und/oder Membrane für medizinische Anwendungen, insbesondere zum Vernähen von Wunden und/oder als Gerüste bzw. als Abdeckung für künstliche Organe, eingesetzt werden.
36. Verwendung nach Anspruch 35, wobei die Folien und/oder Membrane als Anheftungsflächen für kultivierte Zellen und/oder zur Filterung verwendet werden.

Hierzu 21 Seite(n) Zeichnungen

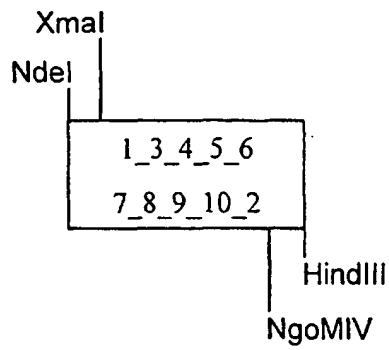
Nr. des Oligos	Sequenz in 5' to 3' Richtung
1	TATGAGCGCTCCCGGGCAGGGT
2	AGCTTTTAGGTACCAATATTAATCTGGCCGGCTCCACC
3	TATGGTCTGGGG
4	GGCCAGGGTGCTGGCCAA
5	GGTGCAGGAGCWGCWGCWGCWGCWGTGCAGGTGGA
6	GCCGGCCAGATTAATATTGGTACCTAAA
7	CTGCCCCGGGAGCGCTCA
8	ACCACCATAACCTCC
9	AGCACCTGGCCCCCAG
10	TGCAGCWGCWGCWGCWGCWGTCTCCTGCACCTTGGCC
11	TATGAGATCTGGCCAAGGAGGT
12	TTGGCCAGATCTCA
13	AGTCAGGGTGCTGGTCGTGGAGGCCAA
14	TCCACGACCAGCACCTGACTCCCCAG
15	AGTCAGGGCGCTGGTCGTGGGGGACTGGGTGGCCAA
16	ACCCAGTCCCCACGACCAGCGCCCTGACTCCCCAG
17	CTGGGAGGGCAGGGAGCGGGCCAA
18	CGCTCCCTGCCCTCCAGACCTCC

Abb. 1

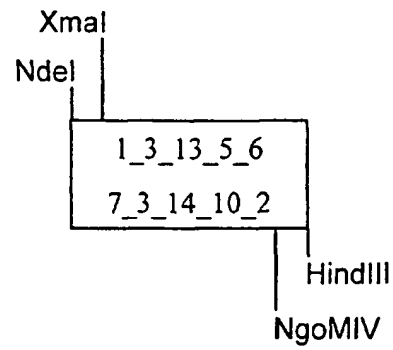
Protein	Anordnung der Genkassetten aus Abbildung 2
SD1	G_D_B_C
SM12	G_D_C_B_C_B_B_G_D_B_C
SO1	H_B_C_B_C_G_D_C_G_D_C_B_C_B_B_G_D_B_C
SO1SO1	H_B_C_B_C_G_D_C_G_D_C_B_C_B_B_G_D_B_C
	H_B_C_B_C_G_D_C_G_D_C_B_C_B_B_G_D_B_C
FA1	G_D_C_B_C_B_B_Fibroin

Abb. 3

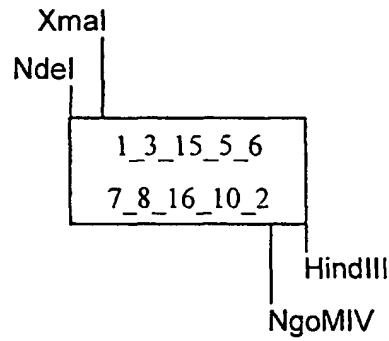
## Abbildung 2



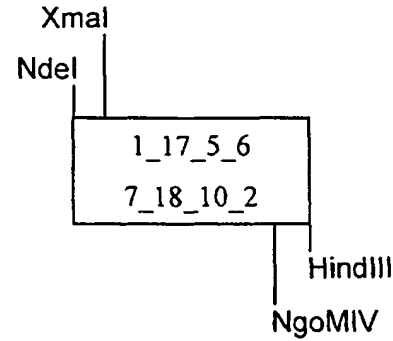
A



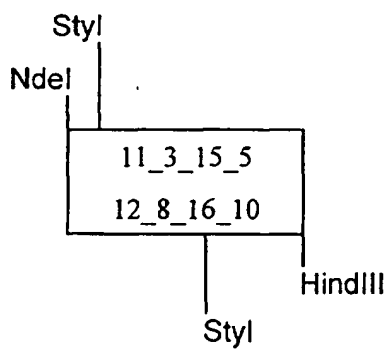
B



C

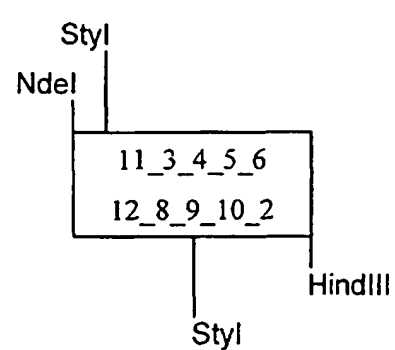


D



E

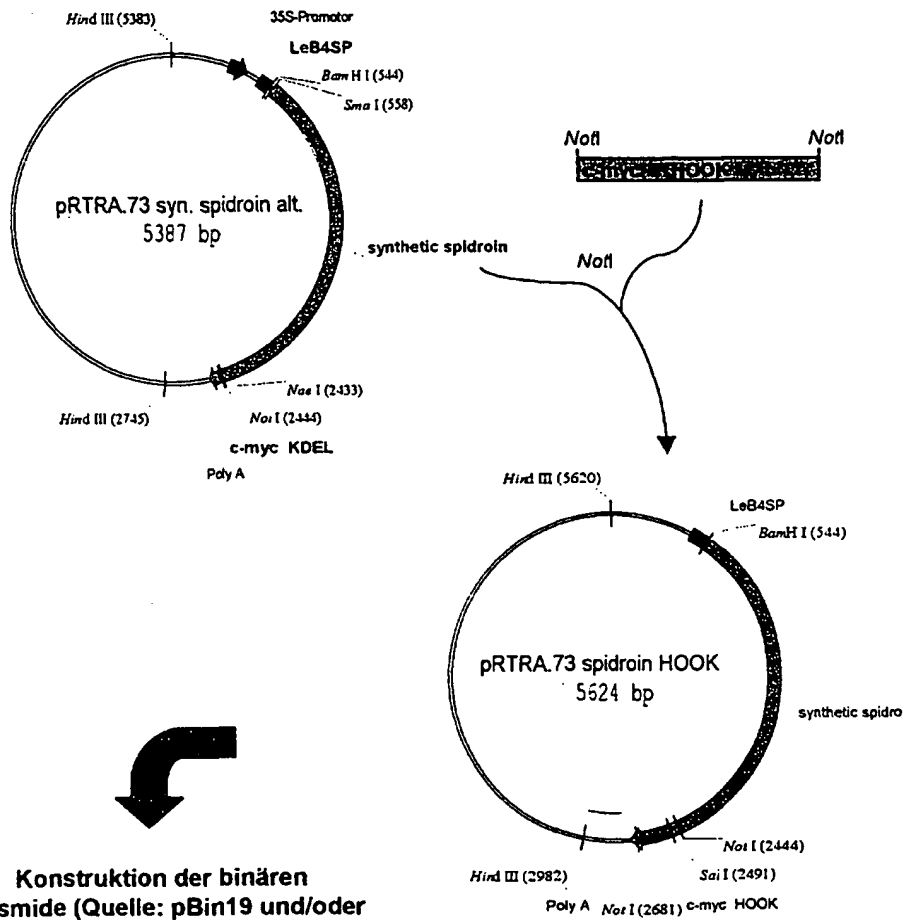
$$G=A+E$$



F

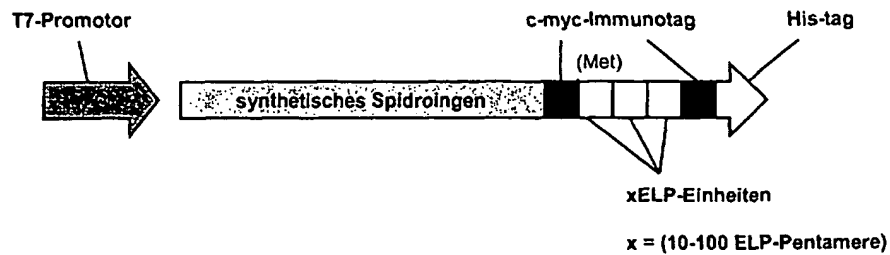
$$H=A+F$$

Abb. 4

 6. Klonierung des Gens der Transmembrandomäne von HOOK mit *NotI* aus (pRT-HOOK) in (pRTRA.73 syn. spidroin)


**Konstruktion der binären  
Plasmide (Quelle: pBin19 und/oder  
pGSGluc1) mit Spidroin-HOOK  
Expressions-kassetten durch  
Transfer von *Hind III*-Fragmenten**

## Spidroin-ELP-Expressionskassetten für *Escherichia coli*



## Spidroin-ELP-Expressionskassetten für transgene Pflanzen

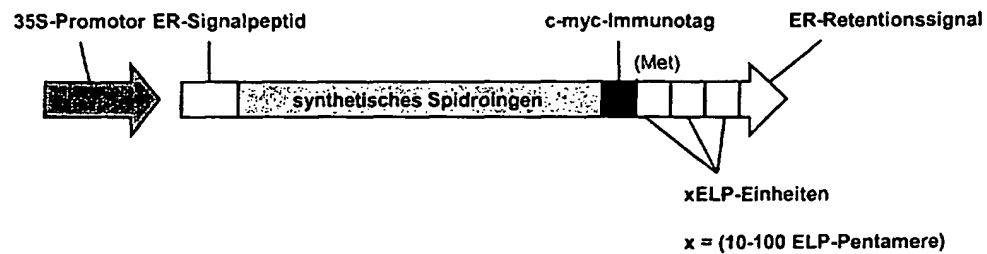
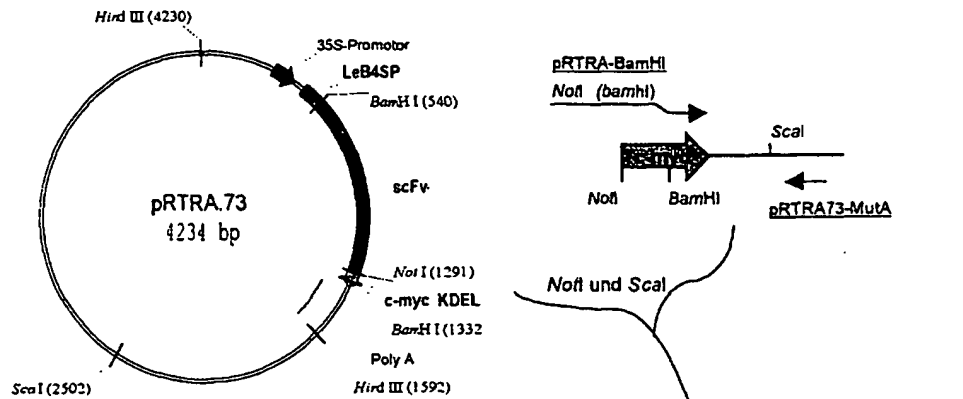


Abb. 5



Abb. 6

# 1. Veränderung einer Base in der Erkennungssequenz von BamHI (Pos. 1332) durch gezielte Mutagenese



# 2. Klonierung synthetischer Spidroingene als PCR-Produkt über BamHI und SmaI-Erkennungssequenzen (Matrize: synthetische Spidroingene aus p9609 oder p9905 nach Ligierung von Oligonukleotiden [ON1 + ON2] als Primerbindungsstellen) in den Vektor (pRTRA.73 BamHIΔ)

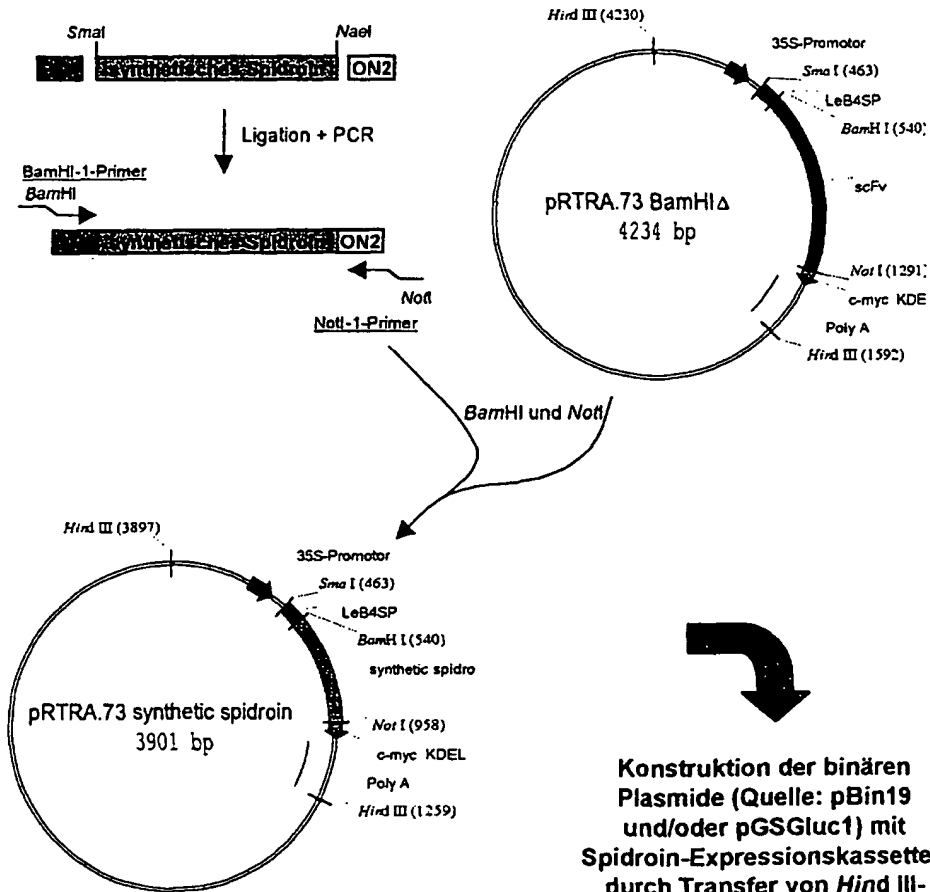


Abb. 7

3. Vorbereiten von (pRTRA.73 BamHI $\Delta$ ) auf die direkte Klonierung der synthetischen Spidroingene aus p9905 oder p9609 - Aufheben der SmaI-Erkennungssequenz (Pos. 463)

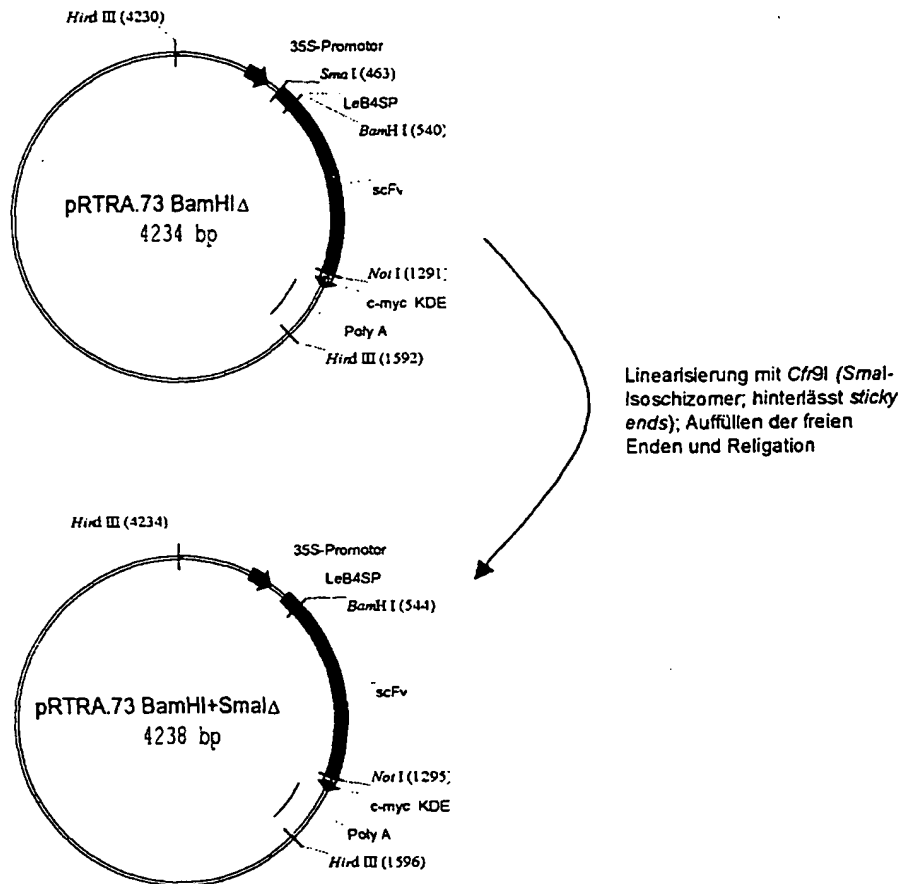
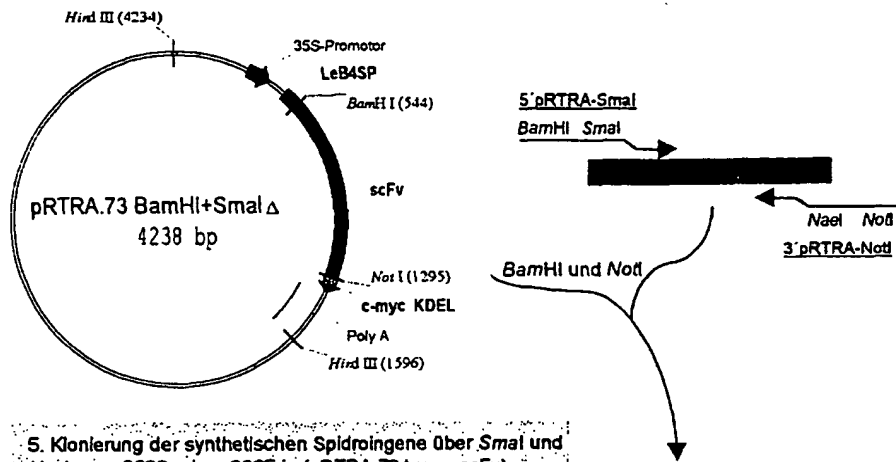
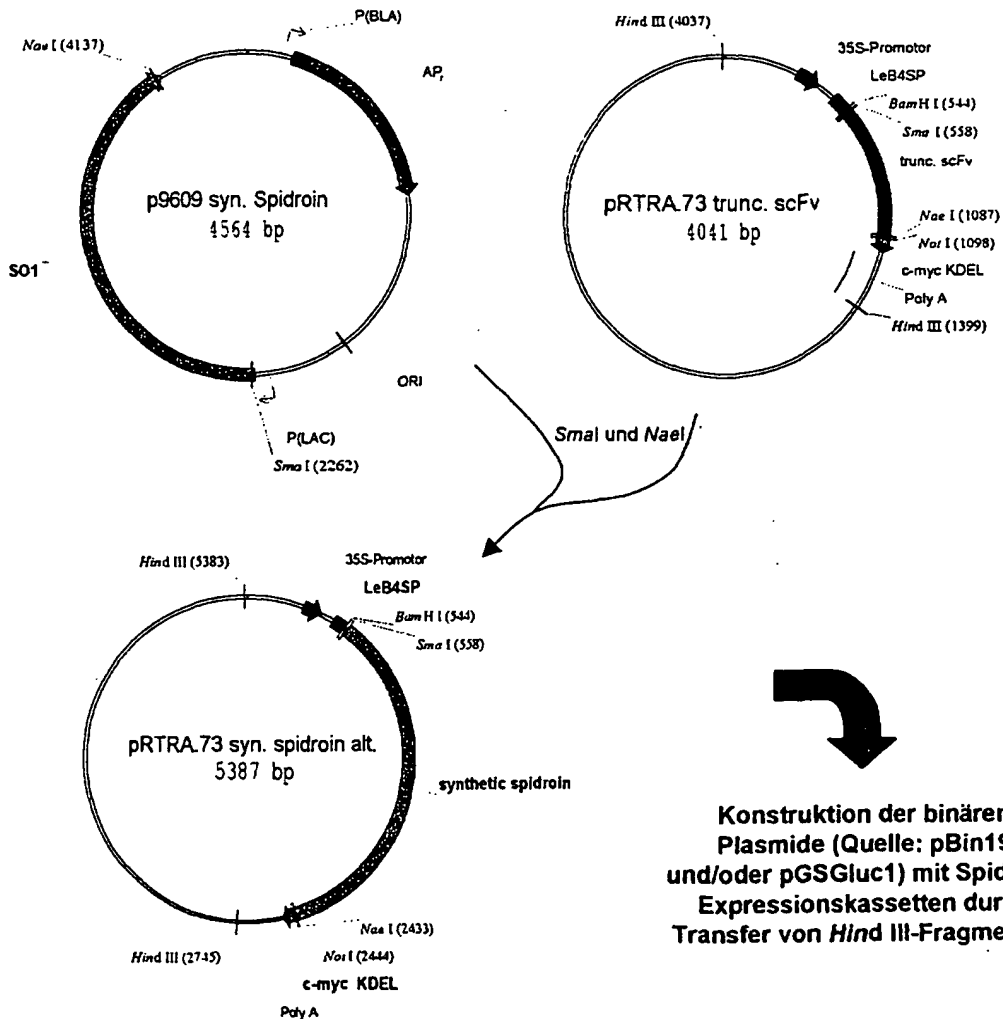


Abb. 8

**4. Einführung der Restriktionserkennungssequenzen von *Sma*I und *Nae*I in den Vektor (pRTRA.73 BamHI+SmaI Δ) für die Klonierung synthetischer Spidroingene**



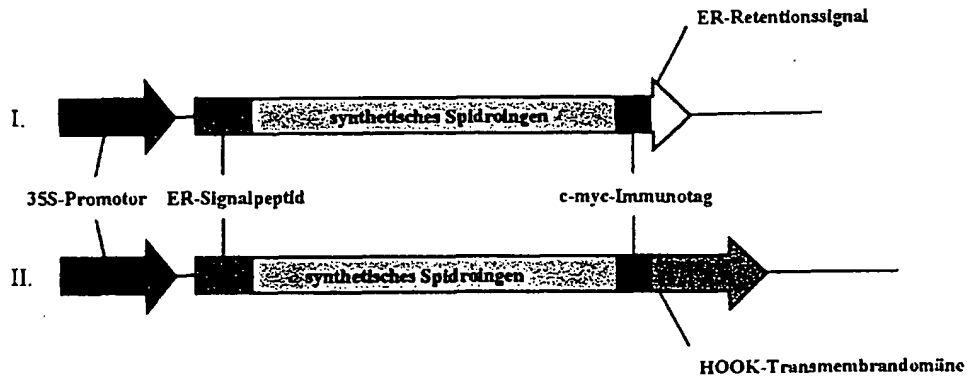
**5. Klonierung der synthetischen Spidroingene über *Sma*I und *Nae*I aus p9609 oder p9905 in (pRTRA.73 trunc. scFv)**



**Konstruktion der binären Plasmide (Quelle: pBin19 und/oder pGSGluc1) mit Spidroin-Expressionskassetten durch Transfer von *Hind* III-Fragmenten**

Abb. 9

**Generelle Darstellung der Herstellung transgener Spidroin-produzierender Pflanzen**

 1. Konstruktion geeigneter pflanzenkompatibler Spidroinexpressionskassetten in *Escherichia coli*


a. Überprüfung der Expressionskassetten durch Sequenzierung und/oder Testexpression im pet-System in *Escherichia coli*

 2. Einbringen der Expressionskassetten in *shuttle*-Vektoren (pBIN19; pGSGluc) in *Escherichia coli*

 3. Transformation der *shuttle*-Vektoren in *Agrobacterium tumefaciens*-Stämme

a. Überprüfung der Klone durch Southern-Blot-Analyse

## 4. Herstellung transgener Tabak- und Kartoffelpflanzen

5. Analyse der transgenen Tabak- und Kartoffelpflanzen auf Spidroinproteineexpression durch immunochemische Detektion eines c-myc-Tags (Western-Analyse)

Abb. 10a

### LeB4-Signalpeptid

MASKPFLSL

LSLSLLFT

STCLAGSQL? => GQGGYGG LGGQAGQG GYGG LGGQAGQG AGA AAA AGGAGQ GGYGG LG

GLGGYGGQGAGGAAAAAGAGQGGRGAGQS

SQGAGRGGGGGAGAAAAAGGAGQGGYGGLG

GLGGYGGQGAGGAAAAAGAGQGGRGAGQS

SQGAGRGG LGGQCAGA-AAA-AGGACQGGYGG LGGQGAGQGGYGG LG

G L G G Y G G Q G A G G A A A A A G A G Q G A G Q G G L G G Q G A G G A A A A A G A G Q G G L G G R G A G Q S  
 G L G G Y G G Q G A G G A A A A A G A G Q G A G Q G G L G G Q G A G G A A A A A G A G Q G G L G G R G A G Q S

5 SQGAGRGGQGAGAAAAAGGAGQGGYGGTGGQGAGQGGYGGTG

GLGGYGGGQAGGA<sup>1</sup>AAAAAGAGQAGGGLGGQAGGA<sup>2</sup>AAAAAGAGQGGLGGRGAGQS

SQGAGRGGGCGGQGA GA A A A AGGAGQGG?GGLG  
61 GGXGGGGGCGG A A A A GCGGGGCGGG

G L G G Y . G G Q G A G G A A A A A G A G Q G G R G A G C S  
G G

SQGAGRGG LGGQGAGAAAAAGGAGQGG?GGTG  
-GC-CCXCGCGC?CG?A??A?CG?CGCG?CG?CG?

GGGGYGGQGAGGAAAAAGAGQGGAGAGQS  
GGGGYGGQGAGGAAAAAGAGQGGAGAGQS

[illegible][illegible]

SQGAGRGGCGAGA--AAAGGAGUGGYGGG

MAAQQGAGGA--AAAGAGQGGGGAGAGS

FOUR SEEDINGS = 32 FL.

c-myc-Immunotag	ER-Retentionssignal
+	+
+	-
-	+
-	-

Abb. 10b

### LeB4-Signalpeptid

MASKPFLSL

LSLSLLFT

STCLAGSQL ⇒ GQGGYGGGGGQAGQGGYGGGSGQAGRGGGGGQAGAAAAAGGAGQGGG

GLGGYGGQGAGGAA-~~AA~~AGAGQGAGQG

5 SQGAG?GGGGGAGAAAAAGGAG

AGAGAGAGSAGAGAGYGVGAGAGYGAGYGVGAGAGYGAGAGSGAGAGSGAGAGSGAGAGSGAGAGAGSGAGAGSGS

GTGSSGFGPYVANGGYSGYEYAWSSKSDFFETAGQAAA

EQUILIBRIUMS  $\Rightarrow$  KDEL.

c-myc-Immunotag	ER-Retentionssignal
+	+
+	-
-	+
-	-

Abb. 11

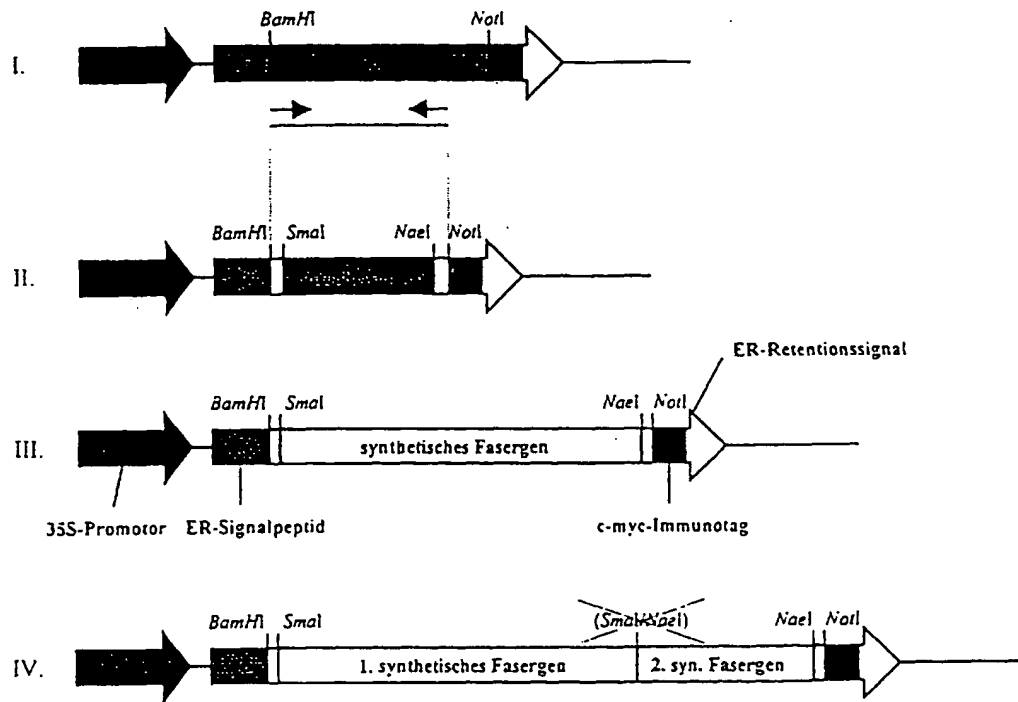
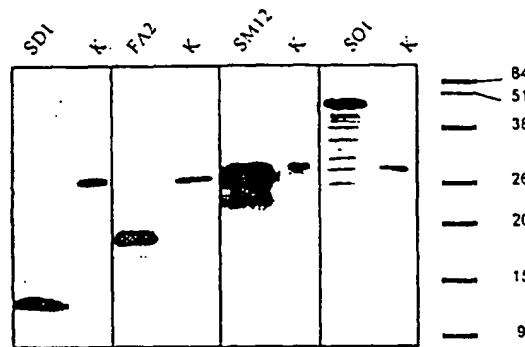
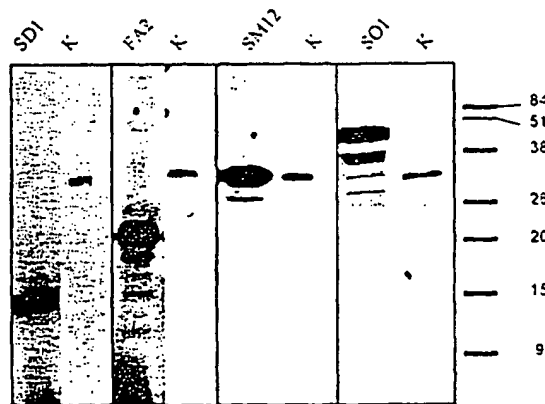


Abb. 12

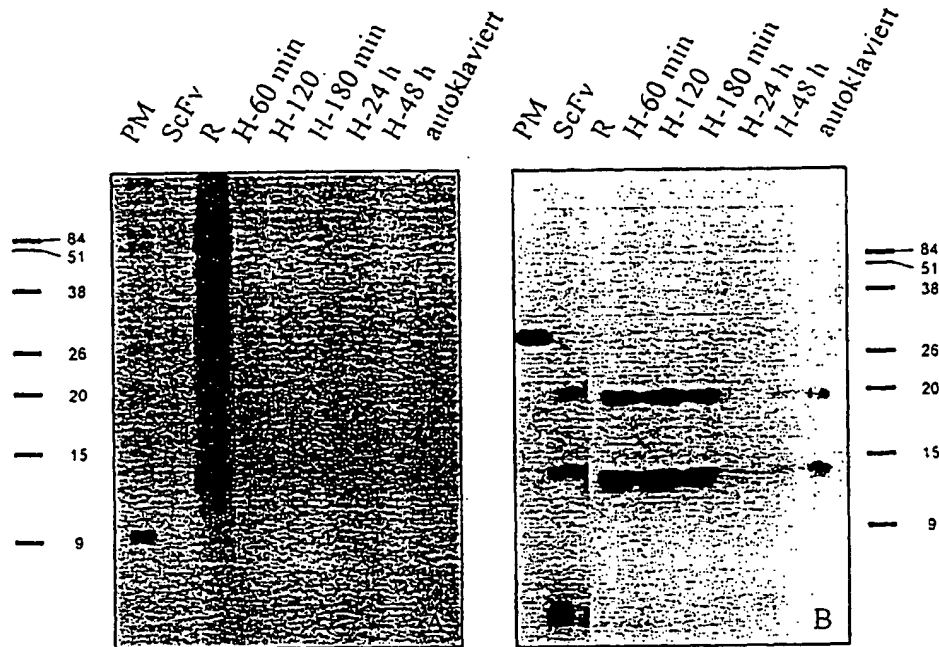


a)



b)

Abb. 13





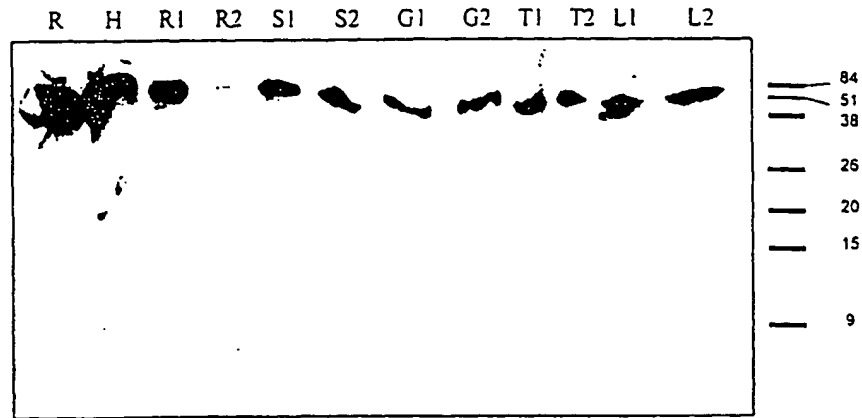


Abb. 14

Abb. 15

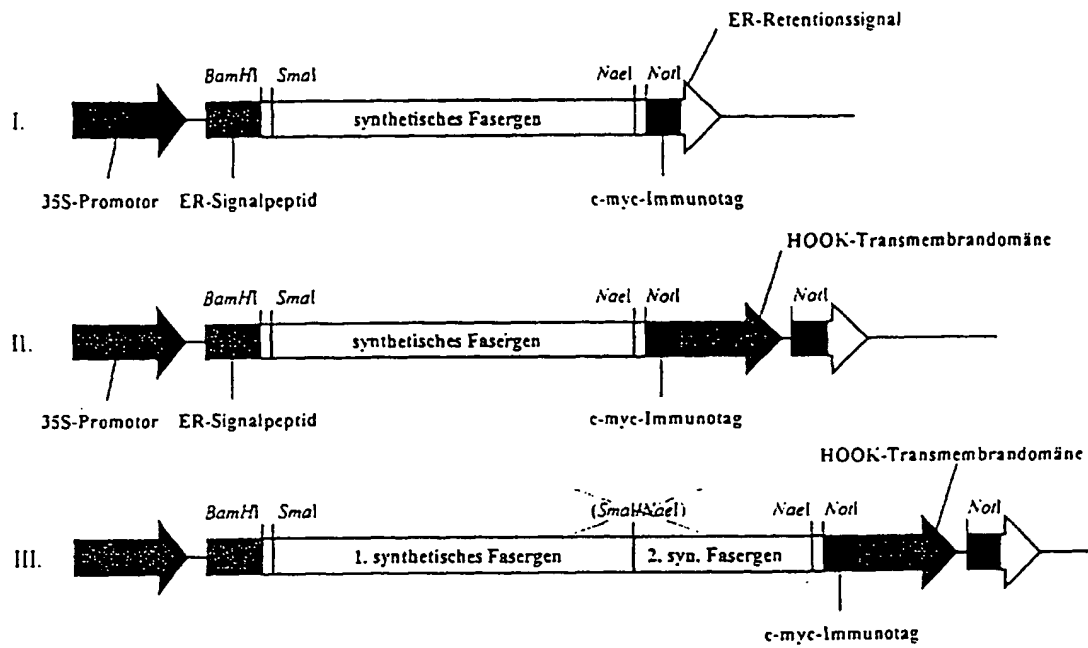
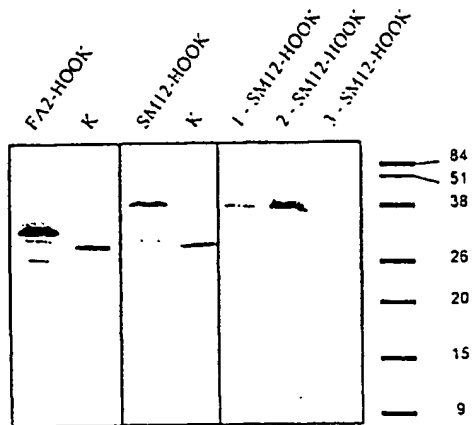


Abb. 16



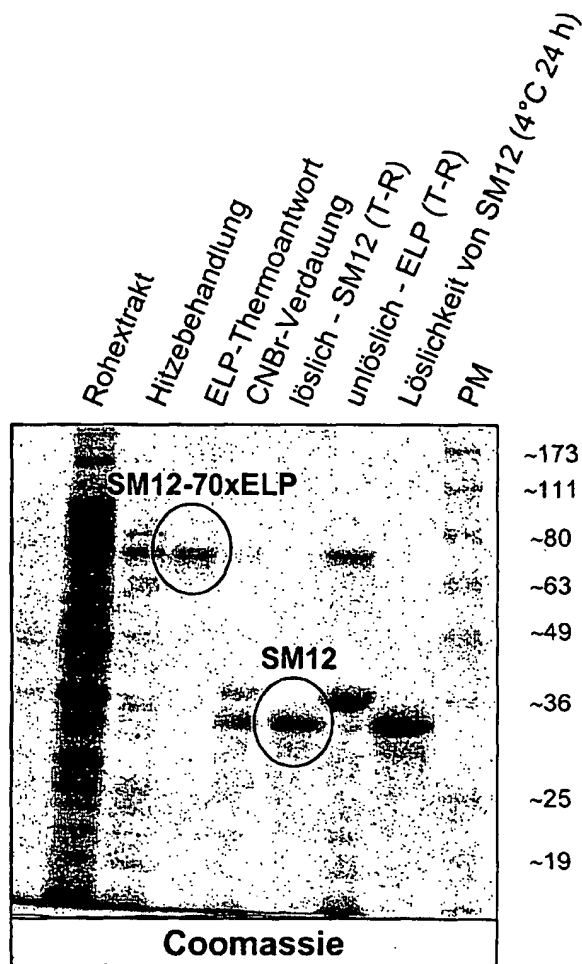
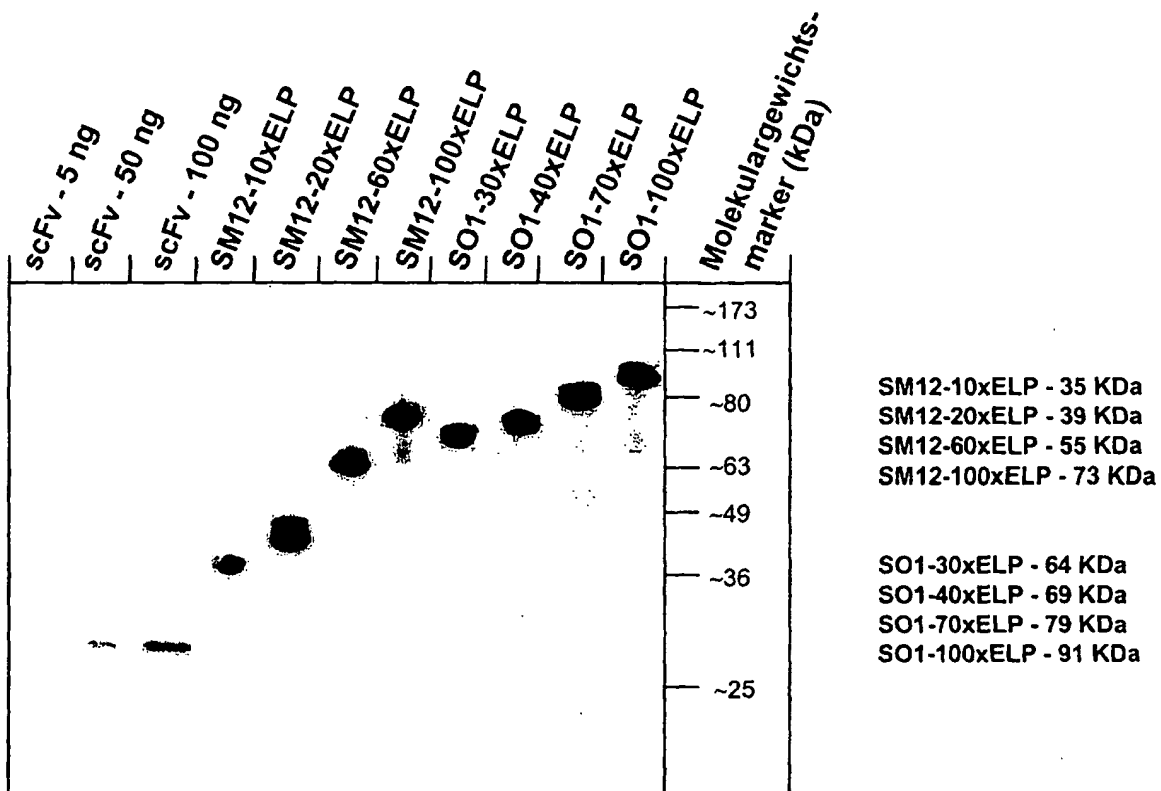


Abb. 17

Abb. 18



**Abbildung 19****DNA-Sequenz von SM12-70xELP als Pflanzenexpressionskassette**

atggcttccaaaccttttctatcttttgcctttcactttccttgcttctctttacaagcacatg  
tttagcaggatcccagttacccgggcagggaggttatgggtggtctggggggccaggggtgctg  
gccaaggaggttatgggtggtctggggagtcagggcgctggtcgtgggggactgggtggccaa  
gggtgcaggagctgctgctgcagctgcaggtggagccgggcagggaggtctgggagggcaggg  
agcgggccaaggtgcaggagcagctgcagcagctgcaggtggagccgggcagggaggttatg  
gtggtctggggagtcagggcgctggtcgtgggggactgggtggccaaggtgcaggagcagct  
gcagctgctgcaggtggagccgggcagggaggttatgggtggtctggggagtcaggggtgctgg  
tcgtggaggccaaggtgcaggagctgcagcagcagctgcaggtggagccgggcagggaggtt  
atgggtggtctggggagtcagggcgctggtcgtgggggactgggtggccaaggtgcaggagca  
gctgcagctgctgcaggtggagccgggcagggaggttatgggtggtctggggagtcaggggtgc  
tggtcgtggaggccaaggtgcaggagctgcagcagcagctgcaggtggagccgggcagggag  
gttatgggtggtctggggagtcaggggtgctggtcgtggaggccaaggtgcaggagctgcagca  
gcagctgcaggtggagccgggcagggaggttatgggtggtctggggggccaggggtgctggcca  
aggaggttatgggtggtctggggagtcagggcgctggtcgtgggggactgggtggccaaggtg  
caggagctgctgctgcagctgcaggtggagccgggcagggaggtctgggagggcagggagcg  
ggccaaggtgcaggagcagctgcagcagctgcaggtggagccgggcagggaggttatgggtgg  
tctggggagtcaggggtgctggtcgtggaggccaaggtgcaggagctgcagcagcagctgcag  
gtggagccgggcagggaggttatgggtggtctggggagtcagggcgctggtcgtgggggactg  
ggtggccaaggtgcaggagcagctgcagctgctgcaggtggagccggcggaagcgccgc  
agaacaaaaactcatctcagaagaggatctgaatggggccgtcgagatggggcaacggcgctgg  
gtgttccgggcgtgggtgttccgggtggcggtgtgcccgggcaggtgttccgtggttaggt  
gtgccgggtgttggtgtgccgggtgttggtgtaccaggtggcggtgttccgggtgcaggcgt  
tccgggtggcggtgtgccgggcgtgggtgttccgggcgtgggtgttccgggtggcggtgtgc  
cgggcgcaggtgttccgtggttaggtgtgccgggtgttggtgtgccgggtgttggtgtacca  
ggtggcggtgttccgggtgcaggcgttccgggtggcggtgtgccgggcgtgggtgttccggg  
cgtgggtgttccgggtggcggtgtgccgggcgcaggtgttccgtggttaggtgtgccgggtg  
ttggtgtgccgggtgttggtgtaccaggtggcggtgttccgggtgcaggcgttccgggtggc  
ggtgtgccgggcgtgggtgttccgggcgtgggtgttccgggtggcggtgtgccgggcgcagg  
tgttccgtggttaggtgtgccgggtgttggtgtgccgggtgttggtgtaccaggtggcggtg  
ttccgggtgcaggcgttccgggtggcggtgtgccgggcgtgggtgttccgggcgtgggtgtt  
ccgggtggcggtgtgccgggcgcaggtgttccgtggttaggtgtgccgggtgttggtgtgcc  
gggtgttggtgtaccaggtggcggtgttccgggtgcaggcgttccgggtggcggtgtgccgg  
gcgtgggtgttccgggcgtgggtgttccgggtggcggtgtgccgggcgcaggtgttccgtggt  
gtaggtgtgccgggtgttggtgtgccgggtgttggtgtaccaggtggcggtgttccgggtgc  
aggcgttccgggtggcggtgtgccgggcgtgggtgttccgggcgtgggtgttccgggtggcg  
gtgtgccgggcgcaggtgttccgtggttaggtgtgccgggtgttggtgtgccgggtgttggt  
gtaccaggtggcggtgttccgggtgcaggcgttccgggtggcggtgtgccgggcgggctggc  
ggccgcagaacccaaagacgaactctag

**Abbildung 20**

**Proteinsequenz von SM12-70xELP aus pflanzlicher Expression (SM12, c-myc-Tag, 70xELP, KDEL – jeweils durch Absatz gekennzeichnet)**

MASKPFLSLLSLSLLLFTSTCLAGSQLPGQGGYGGLGGQAGQGGYGGLGSQAGRGGLGGQ  
GAGAAAAAGGAGQGGGLGGQAGQAGAGAAAAAGGAGQGGYGGLGSQAGRGGLGGQAGAA  
AAAAGGAGQGGYGGLGSQAGRGQGAGAAAAAGGAGQGGYGGLGSQAGRGGLGGQAGAA  
AAAAAGGAGQGGYGGLGSQAGRGQGAGAAAAAGGAGQGGYGGLGSQAGRGQGAGAAA  
AAAGGAGQGGYGGLGGQAGQGGYGGLGSQAGRGGLGGQAGAAAAAGGAGQGGGLGGQGA  
GQGAGAAAAAGGAGQGGYGGLGSQAGRGQGAGAAAAAGGAGQGGYGGLGSQAGRGGL  
GGQAGAAAAAGGAGGQAAA

EQKLISEEDLNGAVE

MGHGVGPVGVPGGGVPGAGVPGVGPVGVPVGVPGGGVPGAGVPGGVPGVGPVGVP  
PGGVPGAGVPGVGPVGVPVGVPGGGVPGAGVPGGVPGVGPVGVPGGGVPGAGVPG  
VGVPVGVPVGVPGGGVPGAGVPGGVPGVGPVGVPGGGVPGAGVPGVGPVGVPVGVP  
VPGGVPGAGVPGGVPGVGPVGVPGGGVPGAGVPGVGPVGVPVGVPGGGVPGAGVPG  
GGVPGVGPVGVPGGGVPGAGVPGVGPVGVPVGVPGGGVPGAGVPGGVPGVGPVGVP  
GVPGGGVPGAGVPGVGPVGVPVGVPGGGVPGAGVPGGVPGGLAAAE

KDEL

**Abbildung 21****DNA-Sequenz von SM12-70xELP als Expressionskassette für *E. coli***

atggctagcatgactgggtggacagcaaattgggtcgcggtatcccagttaccggggcagggagg  
ttatgggtgggtctggggggccaggggtgctggccaaggaggttatgggtgggtctggggagtcagg  
gcgctgggtcgtgggggactgggtggccaagggtgcaggagctgctgctgcagctgcaggtgga  
gccgggcagggaggtctgggagggcagggagcgggccaagggtgcaggagcagctgcagcagc  
tgcaggtggagccgggcagggaggttatgggtgggtctggggagtcagggcgctgggtcgtgggg  
gactgggtggccaagggtgcaggagcagctgcagctgctgcaggtggagccgggcagggaggt  
tatgggtgggtctggggagtcaggggtgctgggtcgtggaggccaagggtgcaggagctgcagcagc  
agctgcaggtggagccgggcagggaggttatgggtgggtctggggagtcagggcgctgggtcgtg  
ggggactgggtggccaagggtgcaggagcagctgcagctgctgcaggtggagccgggcagggga  
ggttatgggtgggtctggggagtcaggggtgctgggtcgtggaggccaagggtgcaggagctgcagc  
agcagctgcaggtggagccgggcagggaggttatgggtgggtctggggagtcaggggtgctgggtc  
gtggaggccaagggtgcaggagctgcagcagcagctgcaggtggagccgggcagggaggttat  
gggtgggtctggggggccaggggtgctggccaaggaggttatgggtgggtctggggagtcagggcgct  
tgggtcgtgggggactgggtggccaagggtgcaggagctgctgctgcagctgcaggtggagccg  
ggcagggaggtctgggagggcagggagcgggccaagggtgcaggagcagctgcagcagctgca  
gggtggagccgggcagggaggttatgggtgggtctggggagtcaggggtgctgggtcgtggaggcca  
agggtgcaggagctgcagcagcagctgcaggtggagccgggcagggaggttatgggtgggtctgg  
ggagtcagggcgctgggtcgtgggggactgggtggccaagggtgcaggagcagctgcagctgct  
gcaggtggagccgggcggacaagcggccgcagaaacaaaactcatctcagaagaggatctgaa  
tggggcgctcgagatggggccacggcggtgggtgttccgggcgtgggtgttccgggtggcggtg  
tgccgggcgcaggtgttccctgggtgtagggtgtgccgggtgttgggtgtgccgggtgttgggtgt  
ccaggtggcggtgttccgggtgcaggcggttccgggtggcggtgtgccgggcgtgggtgttcc  
ggcggtgggtgttccgggtggcggtgtgccgggcgcaggtgttccctgggtgtagggtgtgccgg  
gtgttgggtgtgccgggtgttgggtgtaccaggtggcggtgttccgggtgcaggcggttccgggt  
ggcggtgtgccgggcgtgggtgttccgggcgtgggtgttccgggtggcggtgtgccgggcgc  
agggtgttccctgggtgtagggtgtgccgggtgttgggtgtgccgggtgttgggtgtaccaggtggcg  
gtgttccgggtgcaggcggttccgggtggcggtgtgccgggcgtgggtgttccgggcgtgggt  
gttccgggtggcggtgtgccgggcgcaggtgttccctgggtgtagggtgtgccgggtgttgggtgt  
gccgggtgttgggtgtaccaggtggcggtgttccgggtgcaggcggttccgggtggcggtgtgc  
cgggcgtgggtgttccgggcgtgggtgttccgggtggcggtgtgccgggcgcaggtgttccct  
gggtgtagggtgtgccgggtgttgggtgtgccgggtgttgggtgtaccaggtggcggtgttccggg  
tgcaggcggttccgggtggcggtgtgccgggcgtgggtgttccgggcgtgggtgttccgggtg  
gcgggtgtgccgggcgcaggtgttccctgggtgtagggtgtgccgggtgttgggtgtgccgggtgtt  
gggtgtaccaggtggcggtgttccgggtgcaggcggttccgggtggcggtgtgccgggcgtggg  
tgttccgggcgtgggtgttccgggtggcggtgtgccgggcgcaggtgttccctgggtgtagggtg  
tgccgggtgttgggtgtgccgggtgttgggtgtaccaggtggcggtgttccgggtgcaggcggtt  
ccgggtggcggtgtgccgggcgggctggcgggcgcagaaacaaaactcatctcagaagagga  
tctgaatggggcgctcgagcaccaccaccaccactga



**Abbildung 22**

**Proteinsequenz von SM12-70xELP aus bakterieller Expression (SM12, c-myc-Tag, 70xELP, c-myc-Tag, HisTag – jeweils durch Absatz gekennzeichnet)**

MASMTGGQQMGRGSQLPGQGGYGGLGGQGAGQGGYGGLGSQGAGRGGGLGGQ  
GAGAAAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGR  
GLGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGY  
GGLGSQGAGRGGGLGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAGAAA  
AAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQ  
GGYGGLGSQGAGRGGGLGGQGAGAAAAAAGGAGQGGLGGQGAGQGAGAAAAAA  
GGAGQGGYGGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGG  
GGQGAGAAAAAAGGAGGQAAA

EQKLISEEDLNGAVE

MGHGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGV  
PGVGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPG  
VGVPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVG  
VPGVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGV  
GVGVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGV  
GVPGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGVGVPGVGV  
PGGGVPGAGVPGVGVPGVGVPGVGVPGGGVPGAGVPGGGVPGGLAAA

EQKLISEEDLNGAVEHHHHH.